

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:10:59 ; Search time 128.03 Seconds
(without alignments)
985.003 Million cell updates/sec

Title: US-09-235-416-1
Perfect score: 4030
Sequence: 1 MSGGNIKVVVRFPNARE.....ELRQQQAQMEALTKAQEF 784

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	4030	100.0	784	1	PCT-US99-01355-1	Sequence 1, Appli
2	4030	100.0	784	16	US-09-235-416-1	Sequence 1, Appli
3	4030	100.0	784	20	US-09-654-850-1	Sequence 1, Appli
4	2054	51.0	533	18	US-09-417-507-37298	Sequence 37298, A
5	1691	42.0	1153	21	US-09-718-563-2	Sequence 2, Appli
6	1691	42.0	1153	21	US-09-718-804-2	Sequence 2, Appli
7	1691	42.0	1153	21	US-09-718-842-2	Sequence 2, Appli
8	1680.5	41.7	1770	21	US-09-718-563-6	Sequence 6, Appli
9	1680.5	41.7	1770	21	US-09-718-804-6	Sequence 6, Appli
10	1680.5	41.7	1770	21	US-09-718-842-6	Sequence 6, Appli

Hines
09/235416
seq. inter

11	1669.5	41.4	1793	23	US-60-258-275-444	Sequence 444, App
12	1668.5	41.4	1816	21	US-09-718-563-8	Sequence 8, Appli
13	1668.5	41.4	1816	21	US-09-718-804-8	Sequence 8, Appli
14	1668.5	41.4	1816	21	US-09-718-842-8	Sequence 8, Appli
15	1668.5	41.4	1816	23	US-60-130-217-6	Sequence 6, Appli
16	1660.5	41.2	893	18	US-09-488-725A-3179	Sequence 3179, Ap
17	1658	41.1	1103	18	US-09-467-946-1	Sequence 1, Appli
18	1635	40.6	1773	23	US-60-167-217-18615	Sequence 18615, A
19	1635	40.6	1773	23	US-60-173-464-15257	Sequence 15257, A
20	1635	40.6	1773	23	US-60-191-637-18572	Sequence 18572, A
21	1635	40.6	1773	23	US-60-191-681-14693	Sequence 14693, A
22	1431.5	35.5	1362	19	US-09-580-828-2	Sequence 2, Appli
23	1396.5	34.7	1438	23	US-60-161-932-1631	Sequence 1631, Ap
24	1396.5	34.7	1921	23	US-60-167-217-15727	Sequence 15727, A
25	1396.5	34.7	1921	23	US-60-173-464-12890	Sequence 12890, A
26	1396.5	34.7	1921	23	US-60-191-637-15722	Sequence 15722, A
27	1396.5	34.7	1921	23	US-60-191-681-12455	Sequence 12455, A
28	1311.5	32.5	1375	21	US-09-721-683-2	Sequence 2, Appli
29	1311.5	32.5	1375	21	US-09-721-832-2	Sequence 2, Appli
30	1311.5	32.5	1375	21	US-09-722-139-2	Sequence 119, Appli
31	1278.5	31.7	504	1	PCT-US00-09066-115	Sequence 116, App
32	1276	31.7	503	1	PCT-US00-09066-116	Sequence 6751, Ap
33	1251	31.0	421	18	US-09-488-725A-6751	Sequence 4, Appli
34	1133	28.1	348	21	US-09-718-563-4	Sequence 4, Appli
35	1133	28.1	348	21	US-09-718-804-4	Sequence 4, Appli
36	1133	28.1	348	21	US-09-718-842-4	Sequence 4, Appli
37	1129	28.0	354	21	US-09-718-563-10	Sequence 10, Appli
38	1129	28.0	354	21	US-09-718-804-10	Sequence 10, Appli
39	1129	28.0	354	21	US-09-718-842-10	Sequence 10, Appli
40	1128	28.0	1174	23	US-60-167-217-11913	Sequence 11913, A
41	1128	28.0	1174	23	US-60-173-464-9672	Sequence 9672, Ap
42	1128	28.0	1174	23	US-60-191-637-11936	Sequence 11936, A
43	1128	28.0	1174	23	US-60-191-681-9368	Sequence 9368, Ap
44	1072	26.6	352	19	US-09-580-828-4	Sequence 4, Appli
45	1035	25.7	1121	20	US-09-619-049-1524	Sequence 1524, Ap

ALIGNMENTS

RESULT 1
PCT-US99-01355-1
; Sequence 1, Application PC/TUS9901355
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 18557C-000710PC
; CURRENT APPLICATION NUMBER: PCT/US99/01355
; EARLIER FILING DATE: 1998-01-22
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; FEATURE:
; NAME/KEY: DOMAIN

LOCATION: (443)..(601)
OTHER INFORMATION: stalk domain, unc-104 family domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (602)..(784)
OTHER INFORMATION: tail domain
FEATURE:
NAME/KEY: VARIANT
LOCATION: (713)
OTHER INFORMATION: polymorphic variant #1 Val -> Ile
FEATURE:
NAME/KEY: VARIANT
LOCATION: (762)
OTHER INFORMATION: polymorphic variant #2 Asp -> Glu
FEATURE:
NAME/KEY: VARIANT
LOCATION: (774)
OTHER INFORMATION: polymorphic variant #3 Glu -> Asp
PCT-US99-01355-1

Query Match 100.0%; Score 4030; DB 1; Length 784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60
Db 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60
Qy 61 AFAPDRSYWSEFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTSGKSYSMWG 120
Db 61 AFAPDRSYWSEFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTSGKSYSMWG 120
Qy 121 YGKEGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRLLNPSTKGNLKVRE 180
Db 121 YGKEGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRLLNPSTKGNLKVRE 180
Qy 181 HPSTGPYVEDLAKLVVRSFQETENLMDGNKARTVAATNMNETSSRSKSHAVFTLTQKWH 240
Db 181 HPSTGPYVEDLAKLVVRSFQETENLMDGNKARTVAATNMNETSSRSKSHAVFTLTQKWH 240
Qy 241 DEETKMDTEKVAKISLVDLASERATSTGATGARLKEGAENRSLTGLRVTAALADMS 300
Db 241 DEETKMDTEKVAKISLVDLASERATSTGATGARLKEGAENRSLTGLRVTAALADMS 300
Qy 301 GKQKNQLVPRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSLRYADSAKRK 360
Db 301 GKQKNQLVPRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSLRYADSAKRK 360
Qy 361 NHAVVNEDPNARMIRELKEELAQLRSKLSQSGGGGAGGGGPPVEESYPPDTPLEKQIV 420
Db 361 NHAVVNEDPNARMIRELKEELAQLRSKLSQSGGGGAGGGGPPVEESYPPDTPLEKQIV 420
Qy 421 STQOPDATVKKMSKAEIVEQLNQSEKLYRDLNQWEEKLAKTEIHKEREAALEELGSI 480
Db 421 STQOPDATVKKMSKAEIVEQLNQSEKLYRDLNQWEEKLAKTEIHKEREAALEELGSI 480
Qy 481 EKGFGVPHSKEMPHLVNLSDDPLLAELGLVYNIKPGQTRGVGNVNDTQAEIRLNGSKILK 540
Db 481 EKGFGVPHSKEMPHLVNLSDDPLLAELGLVYNIKPGQTRGVGNVNDTQAEIRLNGSKILK 540
Qy 541 EHCTPENVDNVYTVIPNEKAAVMVNGVRIDKPTLRSGYRIILGDFHIFRNNHPEARAE 600
Db 541 EHCTPENVDNVYTVIPNEKAAVMVNGVRIDKPTLRSGYRIILGDFHIFRNNHPEARAE 600
Qy 601 ROEQSLLRHSVTNSOLGSPAPGRHRTLSKAGSDAGDSRSDSPLPHFRGKDSDFYARR 660
Db 601 ROEQSLLRHSVTNSOLGSPAPGRHRTLSKAGSDAGDSRSDSPLPHFRGKDSDFYARR 660
Qy 661 EASAILGLDQKISHLTDELDAFDVQKARAVRRGLVEDNEEDSDOSSFPVRDKYMSN 720
Db 661 EASAILGLDQKISHLTDELDAFDVQKARAVRRGLVEDNEEDSDOSSFPVRDKYMSN 720

Qy 721 GTIDNFSLDTAITMPGTPRSDDDGDALFFGDKKSKODASNVDEVLELROOQAMEALKTA 780
Db 721 GTIDNFSLDTAITMPGTPRSDDDGDALFFGDKKSKODASNVDEVLELROOQAMEALKTA 780
Qy 781 KOEF 784
Db 781 KOEF 784

RESULT 2

US-09-235-416-1
Sequence 1, Application US/09235416A
GENERAL INFORMATION:
APPLICANT: Sakowicz, Roman
APPLICANT: Goldstein, Lawrence S. B.
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
FILE REFERENCE: 18557C-000710US
CURRENT APPLICATION NUMBER: US/09/235,416A
CURRENT FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: WO PCT/US99/01355
EARLIER FILING DATE: 1999-01-22
EARLIER APPLICATION NUMBER: US 60/072,361
EARLIER FILING DATE: 1998-01-23
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 784
TYPE: PRT
ORGANISM: Thermomyces lanuginosus
FEATURE:
OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
OTHER INFORMATION: microtubule motor protein
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(357)
OTHER INFORMATION: kinesin-like microtubule motor domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (358)..(442)
OTHER INFORMATION: neck domain links motor domain to stalk domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (443)..(601)
OTHER INFORMATION: stalk domain, unc-104 family domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (602)..(784)
OTHER INFORMATION: tail domain
US-09-235-416-1

Query Match 100.0%; Score 4030; DB 16; Length 784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60
Db 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60
Qy 61 AFAPDRSYWSEFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTSGKSYSMWG 120
Db 61 AFAPDRSYWSEFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTSGKSYSMWG 120
Qy 121 YGKEGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRLLNPSTKGNLKVRE 180
Db 121 YGKEGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRLLNPSTKGNLKVRE 180
Qy 181 HPSTGPYVEDLAKLVVRSFQETENLMDGNKARTVAATNMNETSSRSKSHAVFTLTQKWH 240
Db 181 HPSTGPYVEDLAKLVVRSFQETENLMDGNKARTVAATNMNETSSRSKSHAVFTLTQKWH 240

```

; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-654-850-1

Query Match      100.0%; Score 4030; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 MSGGNIKVVVRVRFNAREIDRGAKCIVRMEGNOTILTPPGAEEKARKSGKTIMDGP 60
DB 1 MSGGNIKVVVRVRFNAREIDRGAKCIVRMEGNOTILTPPGAEEKARKSGKTIMDGP 60

QY 61 AFADRSYFDFKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGOTSGKSYSMWG 120
DB 61 AFADRSYFDFKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGOTSGKSYSMWG 120

QY 121 YGKEGVIPRICODMFRINELQKDNLCTVEYSYLEINERVROLLNPSTKGNLAYRE 180
DB 121 YGKEGVIPRICODMFRINELQKDNLCTVEYSYLEINERVROLLNPSTKGNLAYRE 180

QY 181 HPSTGYPVEDLAKLVRSFOEINLMDGNKARTVAATNNNETSSSHAVFTLTLOKWH 240
DB 181 HPSTGYPVEDLAKLVRSFOEINLMDGNKARTVAATNNNETSSSHAVFTLTLOKWH 240

QY 241 DEETKMDTEKVAKISLDVLAGESEATSGTCARLKEGAENBSLSTLGRVIAALADMS 300
DB 241 DEETKMDTEKVAKISLDVLAGESEATSGTCARLKEGAENBSLSTLGRVIAALADMS 300

QY 301 GQKKNLQVPRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSLRYADSAKRIK 360
DB 301 GQKKNLQVPRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSLRYADSAKRIK 360

QY 361 NHAVVNDPNARMIRELKEELAQRLSKLOSSGGGGAGGGGPGVEESYPDDTFLKQIV 420
DB 361 NHAVVNDPNARMIRELKEELAQRLSKLOSSGGGGAGGGGPGVEESYPDDTFLKQIV 420

QY 421 STQPDATVYKKSKEAIVEQUNQSEKLYRDLNQTWEKLAETEEIHKREAALEELGISI 480
DB 421 STQPDATVYKKSKEAIVEQUNQSEKLYRDLNQTWEKLAETEEIHKREAALEELGISI 480

QY 481 EKGFGVGYHKSKEPHLVNLSDDLPAECLVYNIKPGOTRVGNVNDTQAEIRLNGSKILK 540
DB 481 EKGFGVGYHKSKEPHLVNLSDDLPAECLVYNIKPGOTRVGNVNDTQAEIRLNGSKILK 540

QY 541 EHCFTFENVNDVTIVPNKAAVWVNGVRIDKPTLRSGYRIILGDPIFRNHPPEEARAE 600
DB 541 EHCFTFENVNDVTIVPNKAAVWVNGVRIDKPTLRSGYRIILGDPIFRNHPPEEARAE 600

QY 601 REQSLRLHSVTNSQLGSPAPGRHRTLTLSKAGSDADGSDSDSPLPHFRGKDSDFYARR 660
DB 601 REQSLRLHSVTNSQLGSPAPGRHRTLTLSKAGSDADGSDSDSPLPHFRGKDSDFYARR 660

QY 661 EASAILGLDQKISHLTDDELDAIFDDVQKARAVRGLVEDNEDSDSSOSPPVYRDYKMS 720
DB 661 EASAILGLDQKISHLTDDELDAIFDDVQKARAVRGLVEDNEDSDSSOSPPVYRDYKMS 720

QY 721 GTIDNFSLDTAITMPTGTPRSDGCDALFPGDKKSKQDASNVVDVEELRQQQAQMEALKTA 780
DB 721 GTIDNFSLDTAITMPTGTPRSDGCDALFPGDKKSKQDASNVVDVEELRQQQAQMEALKTA 780

QY 781 KQEF 784
DB 781 KQEF 784

RESULT 4
US-09-417-507-37298
; Sequence 37298, Application US/09417507
; GENERAL INFORMATION:

```



```

QY 589 FRFNHPEARER-----QEOSLLRHVTSNQ-----LGSP 619
      |||||:|||||
      | : : : |
Db 586 FRFNHPEARAEKTPSAETSEFVDWTFQAQRELLKQIGDMQKQEMEKRLQEMELYYK 645
      | : : : |
QY 620 APGRHRTLSKAGSDADGDSRSDS----- 643
      | : : : ||| ||| |||
Db 646 EKEBADLLEQRULDADSDGSDSKRSCBSWKLITSLREKLPPSKLQTVVKGGLPSS 705
      | : : : ||| ||| |||
QY 644 -----PLPHFR--GKDSDFYARREASATILGLDQKISHLTD-----DELDALFD 686
      | : : : ||| ||| ||| : : : | : : |||
Db 706 GKKREPIKMYQIPQRRRLSKDSKWTISDLKIQAVKEICYEVA-LNDRFHSRQIEALAI 764
      | : : : ||| ||| ||| : : : | : : |||
QY 687 DVQKARAVRGLVEDNEDSDSQSFVPVRDKYMSNGTIDNFSLDTAITMPGTPRSD----- 742
      | : : : ||| ||| ||| : : : | : : |||
Db 765 VKMKELCAMYKKDPNE-RDSWRAV-ARDVWDTVGVGDEKIEDVMATCKGSTVDVDDLKVH 822
      | : : : ||| ||| ||| : : : | : : |||
QY 743 -DGDALFPGDKKSKODASNVDEELRQQQAQMEAL 777
      | : : : ||| ||| ||| : : : | : : |||
Db 823 IDKLEDILQEVYKQNNMKMDEIKVLRNKMLKMEKVL 858
      | : : : ||| ||| ||| : : : | : : |||

RESULT 8
US-09-718-563-6
; Sequence 6, Application US/09718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
SS-09-718-563-6

```

Query Match	41.7%;	Score 1680.5;	DB 21;	Length 1770;
Best Local Similarity	47.3%;	Pred. NO. 1.3e-122;		
Matches 353;	Conservative 124;	Mismatches 175;	Indels 95;	Gaps 14;

Qy	4	GGNTKVVVRVPPNAREIDRGAKCIVVMESNQITLLPPPGAEEBKAARKSKTMDGPKAFA	63
Db	3	GASVKVAVRVRPPNSRSTRESKCIQIQMOGNSIINPKPE-----APKPSF	51
Qy	64	FDRSYWSF-DKNAPNARQEDLFQDLGVPLLDNAFKNYNCIFAYGOTGSGKSYSMGVG	122
Db	52	FDYSYWSHTSPDPCFASQNRVYNDIGKEMLLHAFSGYNVICFAYGOTGAGKSYTMMGKQ	111
Qy	123	KEH--GVIPRICQDMFRINELQOKNMLCTVBVSYLEIYNERNVDLLNPSYKGNLKVRE	180
Db	112	EESQAGIIPQCEELPEKIND-NCNEEMSYSVESYMEIYCERVDLLNPKNKGNLRYRE	170
Qy	181	HPSTGTPVEDLAKLVVRSPOEINLMDGKNKARTVAATNNSTSSKSHAVFTLTUTQKWH	240
Db	171	HPLLGTPVEDLSLAVTSYDIDADLMDAGNKARTVAATNNSTSSRSHAVFTVTFQKKH	230
Qy	241	DEETKMDTEKAKISLVDLAGSERATSTGATGARLKEGAEINRSLTIGRVTAALADWS	300
Db	231	DNETNLSTEKVKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEYSK	290
Qy	301	GKQKNQGLVPRDSVLTWLLKDSIGNSMTAMTAATSPADINFEETLSLRYADSAKRIK	360
Db	291	-KKKKTDFIPYRDSVLTWLLRNLGNSRTAMVAALSPADINWDETLSTLYADRAKQIK	349
Qy	361	NHAVNEDPNARMIRELKEBQALRSKLASSGGGGGGAGSGGPPVEESYPPDTPLEKQ--	418
Db	350	CNAVINDPNKAVIRELKEBVTIRKOLLQAQGLGDIIDTSMGSLT--SSPSSCSSLSSQVG	407

```

Qy 419 ---IVSTQDPADTVKMKSAEIVELQNLNQSEKLYRDLNQTWEELKATETIHKREAALEE 475
Db 408 LTSVTSIQ--ERIMSTPGGBEAIERLKSESEKIITAEELNETWEEKLRKTEAIRMERERALLAE 465
Qy 476 LGISIEK*-GFGVGYHSEKMPHLVNLSDPPLLAECPLVYNIKPGQTRGVNVNODTQAEIRL 533
Db 466 MGVAIRDEGGTGLVGFSPKTPHLVNLNEDPIMSECLYYIKDGITRVQADAEERRQDIVL 525
Qy 534 NGSKILKEHCTFENV-----DNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHI 588
Db 526 SGAHKEEHCIERSERSNSGEVITVLEPCERSETTYVNGKRVSPQVLRSNGRIIMCKNHV 585
Qy 589 FRFNHPPEARAEQEQSLLRHSTVNSQLSGPAPGRHDRTLSKAGSDADGDSRSDSPHPHF 648
Db 586 FRFNHPQQAAREK-----TPSAET 606
Qy 649 RGDSDFYARRAAASAILGLDQK-----ISHLTDDELALPDD-----687
Db 607 PSEPVDWTFQARELLEK-QGIDMKQEMKEKLOEMEILYKKEEADLLLEQORLDYESKL 665
Qy 688 --VQKRAVRRLGLVEDNEDSDSQSSFP 712
Db 666 QALQKQVETRSLAAETEEEEEEV 692

RESULT 9
US-09-718-804-6
; Sequence 6, Application US/09718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-6

```

Query Match 41.7%; Score 1680.5; DB 21; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.3e-122;
Matches 333; Conservative 124; Mismatches 175; Indels 95; Gaps

[illegible]

```
QY 361 NHAVNEDPNARMIRELKEELAQLRSLKSSGGGGGAGGCGPVESYPDPDPLEKQ-- 418
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 CNAVINEDPNAKLVRELKEEVTRLDLLRAQGLGDIIDTSMGSLT--SSPSSCSLSSQVG 407
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 419 ---IVTSIQDPATVKKMSKAEIVQLNOSKLYRLDNLNQTWEEKLAKTEIHKEREALAE 475
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 408 LTSVTSIQ--ERIMSTPGCEAIERLKESEKIIAELNETWEEKLAKTEAIRMERALLAE 465
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 476 LGISIEK--GFVGPYHSKEMPHLVNLSDDPLLAECVLNKKPGQTRVGNVNDQABRL 533
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 466 MGVAIREDGCTLGVPSPKKTPLHVLNEDPLMSECLLYIKDGTITRVGQADARRQDIVL 525
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 534 NSKILKEHCTFENV-----DNVVTIVPNEKAAMVNVGRIDKPTRLSRGYRIILGDFHI 588
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 526 SGAIKEEHCIFRSERSNSGEVITLPCERSETYVNGKRVSPQVQLRSGNRIMGNKHNH 585
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 589 FRFNHPEARAERQEOSLLRHVSTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPHPF 648
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 586 FRFNHPEARAEREK-----TPSAET 606
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 649 RGKDSWFYARREASAILGLDQK-----ISHLTDELDALFDD----- 687
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 607 PSEPVDWTFQAERLEK-QGIDMKQEMKRLQPMELIYKKEEADLLLEQORLDYESKL 665
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 688 --VOKARAVRGLVEDNEDSDSSQSPF 712
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 666 QALQKQVETRSLAAETTEEEEEEEVP 692
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 10

```
.US-09-718-842-6
; Sequence 6, Application US/09718842
; GENERAL INFORMATION:
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
; .US-09-718-842-6
```

```
Query Match 41.7%; Score 1680.5; DB 21; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.3e-122;
Matches 353; Conservative 124; Mismatches 175; Indels 95; Gaps 14;
```

```
QY 4 GGNIVVVRVPFNAREIDRGAKCIVRMENQOTILTPPPGAEKARKSGKTIMDGPKAFA 63
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 GASVAVAVRVPNRSKTSKSKCIIOMQNSTSIINPKNPKE-----APKSFS 51
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 64 FDRSYNSF-DKNAPYARQEDLFODLGVLPLDNLNFAKYNQICIFAYGQTSGKSYSMGYG 122
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 52 FDYSVWSHSTPDCPCASQNRVYNDICKEMLLHAFEGYVNCIFAYGQTGAGKSYTMGKQ 111
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 123 KEH--GVIPRICODMFRINELQDKNLCTVEVSYLIYNERVROLLNPKSKGNLKVRE 180
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 112 EESQAGIIPOLCEDLFSRIND--NCNEEMSYSVEVSYMEIYCERVRDOLLNPKNGLRVRE 170
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 181 HPSTGPTVEDLAKLVRSFQBIENLMDGNKARTVAATNMNETSSRSHAVFTLTITQKH 240
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 171 HPLLGPVEDLSKIAVTSYTDIADNLNAGNKARTVAATNMNETSSRSHAVFTVPTQKH 230
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 241 DEETKNDTEKAKISLVLAGSERATSTGATGRLKEGAEINSLSTGLRVTAALADMS 300
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 231 DNETNLSTEVKISLVLAGSERADSTGAKTRLKEGANINKSLATTLGKVISALAEVSK 290
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
QY 301 GKOKNQLYPRDSVLTMLKDSLGNSWTAMIAAISPADINFEETLSTLRYADSAKRK 360
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 291 -KKKTYDFIPYHDSVLTMLKRENLGNSKRTAMVAULSPADINTDTLSTLRYADRAQRIK 349
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 361 NHAVNEDPNARMIRELKEELAQLRSLKSSGGGGGAGGCGPVESYPDPDPLEKQ-- 418
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 350 CNAVINEDPNAKLVRELKEEVTRLDLLRAQGLGDIIDTSMGSLT--SSPSSCSLSSQVG 407
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 419 ---IVTSIQDPATVKKMSKAEIVQLNOSKLYRLDNLNQTWEEKLAKTEIHKEREALAE 475
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 408 LTSVTSIQ--ERIMSTPGCEAIERLKESEKIIAELNETWEEKLAKTEAIRMERALLAE 465
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 476 LGISIEK--GFVGPYHSKEMPHLVNLSDDPLLAECVLNKKPGQTRVGNVNDQABRL 533
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 466 MGVAIREDGCTLGVPSPKKTPLHVLNEDPLMSECLLYIKDGTITRVGQADARRQDIVL 525
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 534 NSKILKEHCTFENV-----DNVVTIVPNEKAAMVNVGRIDKPTRLSRGYRIILGDFHI 588
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 526 SGAIKEEHCIFRSERSNSGEVITLPCERSETYVNGKRVSPQVQLRSGNRIMGNKHNH 585
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 589 FRFNHPEARAERQEOSLLRHVSTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPHPF 648
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 586 FRFNHPEARAEREK-----TPSAET 606
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 649 RGKDSWFYARREASAILGLDQK-----ISHLTDELDALFDD----- 687
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 607 PSEPVDWTFQAERLEK-QGIDMKQEMKRLQPMELIYKKEEADLLLEQORLDYESKL 665
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 688 --VOKARAVRGLVEDNEDSDSSQSPF 712
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 666 QALQKQVETRSLAAETTEEEEEEEVP 692
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 11

```
US-60-258-275-444
; Sequence 444, Application US/60258275
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL001026-PROV
; CURRENT APPLICATION NUMBER: US/60/258,275
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 717
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 1793
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-258-275-444
```

```
Query Match 41.4%; Score 1669.5; DB 23; Length 1793;
Best Local Similarity 46.8%; Pred. No. 9.8e-122;
Matches 352; Conservative 125; Mismatches 166; Indels 109; Gaps 15;
```

```
QY 4 GGNIVVVRVPFNAREIDRGAKCIVRMENQOTILTPPPGAEKARKSGKTIMDGPKAFA 63
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 9 GASVAVAVRVPNRSKTSKSKCIIOMQNSTSIINPKNPKE-----PKSFS 57
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 64 FDRSYNSF-DKNAP--NYARQEDLFODLGVLPLDNLNFAKYNQICIFAYGQTSGKSYSMWG 120
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 58 FDYSYNS--HTSPEDINYSQKQVYRDIGEEMLQHAPEGYNVCIFAYGQTGAGKSYTMWG 115
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 YGK--BHGVIPTICQDMFRINELQDKNLCTVEVSYLIYNERVROLLNPKSKGNLKV 178
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 116 KQEKDQGGIIPOLCEDLFSRIND--NMSYSVEVSYMEIYCERVRDOLLNPKNGLRV 174
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 179 REHPSTGPTVEDLAKLVRSFQBIENLMDGNKARTVAATNMNETSSRSHAVFTLTITOK 238
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 175 REHPLIGPYVEDLSKIAVTSYTDIADNLNAGNKARTVAATNMNETSSRSHAVFTITOK 234
      :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
QY 239 WHDEETKMDTEKVAKISLVLDLAGSERATSTGATCARLKEGAENRSLSLTGRVTAALADM 298
Db 235 RHDAETNITTEKVKISLVLDLAGSERADSTGAKTRLKEGANINSLTTLGKVISALAEM 294
QY 299 SSG-----KQKKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINPEETLSLRYA 353
Db 295 DSGPNKNNKKKTDFIPYRDSVLTWLLRENLCGNSRTAMVAALSPADINDETSLRYA 354
QY 354 DSARKRIKHAVNEDPNARMIRELKEAQLRSKLOSGGGGGAGGGGPPVEESYPPDT 413
Db 355 DRAKOIRCAVINEDPNKRLREKLDVETRLDLYAQLG-----DIT 398
QY 414 PLEKQIVSIQOPDATVKKMS-----KABIVQLNQSEKLYRDLAQTW 455
Db 399 DMTNALVGM-SPSSLSALSRAASVSLHERILFAFGSEEAIRLKEKETEIIAELNETW 457
QY 456 EEKLAKEEETHKEREAALEELGIGIEK--GFVGPYHSGKEMPHLVNLSDDPLLAELVNI 513
Db 458 EEKLRRTAETMEREALLAEMGVAMREDGGTGLGVFSPKKTPHLVNLNEDPLMSCLLYI 517
QY 514 KPGQTRVGNQDQAEIRLNGSKILKEHCTFEN-----VDNVVTIIVPNEKAAVMVNGVR 568
Db 518 KDGITRVGREDGERRQDVLGSHFIKEHCVFSDRSRGSEAVVTLPECEGADTVVNGKK 577
QY 569 IDKTRLSRGYRIITLGDHFIHFRPHPEAREARQEQSLRLHSHVNSQLGSPAGRHDTL 628
Db 578 VTSPSILRSGNRIITMGKSHVFRPHPEAREARER----- 612
QY 629 SKAGSDADGDSRSDPLPHFRGKDSDFYARREASATILGDLQKISHLTDDDELDAFDV 688
Db 613 -----TPCAETPAEPVDNAFAQRELLEK--QGIDMK--QMEORLQLEDOY 655
QY 689 OKARAVRGLVEDNEDSDSQSFVPRDKYMSN 720
Db 656 RREREATYILLE--QORLDYESKLEALQKQMDS 686
RESULT 12
US-09-718-563-8
; Sequence 8, Application US/09718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1816
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-8
Query Match 41.4%; Score 1668.5; DB 21; Length 1816;
Best Local Similarity 45.0%; Pred. No. 1.2e-121;
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;
QY 4 GGNKTVVVRVPFNAREIDRGAKCIVRMEGNQTLTTPPPGAEEKARKSGKTMGPKRAFA 63
Db 3 GASVKAVRVRFPNSRETSKESKCIIOGQNSTSIINPKNPKE-----APKSFS 51
QY 64 FDRSYWQSF--DKNAPNARYAQEDLFODLGVPLLDNAFKYNNCIPAYGQTGSKSYMMGYG 122
Db 52 FDIYSHWTSDEPCFASQNRNQLIDKEMILLHAFEGYNNVCFAYGQTGAGKSYTMGKQ 111
QY 123 KEH--GVIPRICQDMFRINELQKDLNLTCTVEYSYLEIINERNVRLNLPSTKGNLKVRE 180
Db 112 EESQAGIIPQICEELFEKIND--NCNEEMSYSEVSYMEIYCERYVRLNLPKNKNLKVRE 170
```

```
QY 181 HPSTGPPYVEDLAKLVVRSFQEIENIMDEGNKARTVAATNMNETSSRSRHAVFTLTUTQKWH 240
Db 171 HPLGLPYVEDLSKLVATSYTDIADLMDAGNKARTVAATNMNETSSRSRHAVFTIVTQKKH 230
QY 241 DEETKMDTEKVAKISLVLDLAGSERATSTGATCARLKEGAENRSLSLTGRVTAALADM-- 298
Db 231 DNETLSTKRVSKISLVLDLAGSERADSTGAKTRLKEGANINKSLTTTIGKVISALAEDVN 290
QY 299 ---SSGQKKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINPEETLSLRYADS 355
Db 291 CTSKSKKKKTDFIPYRDSVLTWLLRENLCGNSRTAMVAALSPADINDETSLRYADR 350
QY 356 AKRKINHAVNEDPNARMIRELKEAQLRSKLOSGGGG-----GGAG----- 399
Db 351 AKQIKCAVINEDPNARMIRELKEAQLRSKLOSGGGG-----GGAG----- 399
QY 400 -----GSGGPPVEESYPPDTPLEKQ-----IVSIQOPDATVKK 431
Db 411 DFONNKIRYLLASENQRPGHFSTASMGSLTSS--PSSCSLSSOVGLTSVTSTIO--ERIMST 467
QY 432 MSKAETVEQLNQSEKLYRDLAQTWEEKLAKEEETHKEREAALEELGIGIEK--GFVGPYH 489
Db 468 PGEEAETRLKESEKIATLNETWEEKLRKTEAIRMEREALLAEMGVAREGGTGLGVFS 527
QY 490 SKEMPHLVNLSDDPLLAELVNIKPGQTRVGNQDQAEIRLNGSKILKEHCTFENV- 548
Db 528 PKKTPHLVNLNEDPLMSCLLYIKDGITRVGQADAEARRQDVLGSAHIKEHCIFRSE 587
QY 549 ---DNVVTIIVPNEKAAVMVNGVRIDKTRLSRGYRIITLGDHFIHFRPHPEAREARQEQ 604
Db 588 SNSGEVITLPECSRSETYVNGKRVQPVQLSRNRIIMGNKHNHVRFPNHPBOAREAREK- 646
QY 605 SLLRHSVTNSQLGSPAPGRHDTLSKAGSDADGDSRSDPLPHFRGKDSDFYARREAS 664
Db 647 -----TPSAETPPEVDNFAQRELLE 668
QY 665 AILGLDQK-----ISHLTDDDELDAFD-----VOKARAVRGLVED 701
Db 669 K-QGIDMKQEMEKRLQEMEILYKKEKEADULLLEQORLDYESKLEALQKQVETRSLAAET 727
QY 702 NEDSDSQSF 712
Db 728 TEEEEEVEP 738
RESULT 13
US-09-718-804-8
; Sequence 8, Application US/09718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1816
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-8
Query Match 41.4%; Score 1668.5; DB 21; Length 1816;
Best Local Similarity 45.0%; Pred. No. 1.2e-121;
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;
QY 4 GGNKTVVVRVPFNAREIDRGAKCIVRMEGNQTLTTPPPGAEEKARKSGKTMGPKRAFA 63
Db 3 GASVKAVRVRFPNSRETSKESKCIIOGQNSTSIINPKNPKE-----APKSFS 51
```

```
QY 64 FDRSYWSP-DKNAPNVAHQEDLFQDLGVPLLDNAFKYNNCFAYGOTGSGKSYSMGYG 122
Db 52 FDSYWSHTSPEDCFASQNRVYNDIGKEMLLHAFEGYVNCIFAYGOTGAGKSYTMGKQ 111
QY 123 KEH--GVIPRTCDQMFRRNLEQDKNICTVEVSYLEIYNRVRDLNPNSTKGNLKVRE 180
Db 112 EESQAGIIPOLCEBELFEKIND-NCNEEMSYSEVSYMEIYCERVDRLLNPKNGNLVRE 170
QY 181 HPGTGPVYEDLAKLVYRSFOEINLMDGKNGKARTVAATNNNETSSRSHAVFTLTITQKH 240
Db 171 HPLGPGVYEDLSKLAVTSYTDIADMDAGNKARTVAATNNNETSSRSHAVFTLVFTQKH 230
QY 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADM-- 298
Db 231 DNETNLSTEKVKISLVLAGSERADSTGAGTRLKEGANINKSLTTLGKVISALAEVDN 290
QY 299 ---SSGKOKKQNLVYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLYADS 355
Db 291 CTSKSKKKKTKDFIPYRDSVLTWLLREMLGNSRTAMVAALS PADINDETLSLRYADR 350
QY 356 AKRIKHAVNEDPNARMIRELKEELAQRLSKLQSSGGG-----GGAG----- 399
Db 351 AKQIKNAVINEDPNARKLVRELKEBVTRLKDLLRAOGLGDIIDPLIDDYSGSGSKYLK 410
QY 400 -----GSGGPVEESYPDPDTLEKQ-----IVSIQOQDATVKK 431
Db 411 DFQNKRYLLASENQRPCHFTASMGSLTSS-PSSCSLSSQVGLTSVTSIQ--ERIMST 467
QY 432 MSKAEIVQNLQNSKLYRLDNTWEEKLAKTEBIHKEREALAEELGTSIRK--GFVGPYH 489
Db 468 PGCEAIERLKESEKIITAEINLNETWEEKLRKTEAIRMERALLAEMGAIVREDGGTLGVFS 527
QY 490 SKEMPHLVNLSDDPLAECLVNIKPGOTRVGNVNDQAEIRLNGSKILKEHCTFENV- 548
Db 528 PKKTPHLVNLNEDPLMSECLLYIKGTRVQGAADARRQDIVLSGAHKEEHCFRSE 587
QY 549 ---DNVVTIVPNEKAAMVNVGRIDKPTRLSRGYRIILGDPHIFRHNHPEARAEQ 604
Db 588 SNSGEVITVLEPCERSETYVNGKRVSPQVQLRSGNRIIMGNHVFERNHPEQARAREK- 646
QY 605 SLLRHSVTNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRCKDSDFYARREAS 664
Db 647 -----TPSAETPSEPVDWTFAQRELE 668
QY 665 AILGLDQK-----ISHLTDDDELALFDD-----YOKARAVRGLVED 701
Db 669 K-QGIDMKQEMKRLQEMELIYKKEEADLLLEQORLDYESKLQALQKQVETRSLAAET 727
QY 702 NEDSDSOSFP 712
Db 728 TEEEBEEVEP 738
```

```
RESULT 14
US-09-718-842-8
; Sequence 8, Application US/09718842
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1816
; TYPE: PRF
; ORGANISM: Human
US-09-718-842-8
```

```
Query Match 41.4%; Score 1668.5; DB 21; Length 1816;
Best Local Similarity 45.0%; Pred. NO. 1.2e-121;
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;

QY 4 GGNIKVVYVRFPNAREIDRGAKCIVRMGNGNOTILTPPGAEEKARKSGKTTMDGPKAFA 63
Db 3 GASVKVAVRVFPNSRETSKESKCIIOGOGNSTSIINPKNPK-----APKSFS 51
QY 64 FDRSYWSP-DKNAPNVAHQEDLFQDLGVPLLDNAFKYNNCFAYGOTGSGKSYSMGYG 122
Db 52 FDSYWSHTSPEDCFASQNRVYNDIGKEMLLHAFEGYVNCIFAYGOTGAGKSYTMGKQ 111
QY 123 KEH--GVIPRTCDQMFRRNLEQDKNICTVEVSYLEIYNRVRDLNPNSTKGNLKVRE 180
Db 112 EESQAGIIPOLCEBELFEKIND-NCNEEMSYSEVSYMEIYCERVDRLLNPKNGNLVRE 170
QY 181 HPGTGPVYEDLAKLVYRSFOEINLMDGKNGKARTVAATNNNETSSRSHAVFTLTITQKH 240
Db 171 HPLGPGVYEDLSKLAVTSYTDIADMDAGNKARTVAATNNNETSSRSHAVFTLVFTQKH 230
QY 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADM-- 298
Db 231 DNETNLSTEKVKISLVLAGSERADSTGAGTRLKEGANINKSLTTLGKVISALAEVDN 290
QY 299 ---SSGKOKKQNLVYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLYADS 355
Db 291 CTSKSKKKKTKDFIPYRDSVLTWLLREMLGNSRTAMVAALS PADINDETLSLRYADR 350
QY 356 AKRIKHAVNEDPNARMIRELKEELAQRLSKLQSSGGG-----GGAG----- 399
Db 351 AKQIKNAVINEDPNARKLVRELKEBVTRLKDLLRAOGLGDIIDPLIDDYSGSGSKYLK 410
QY 400 -----GSGGPVEESYPDPDTLEKQ-----IVSIQOQDATVKK 431
Db 411 DFQNKRYLLASENQRPCHFTASMGSLTSS-PSSCSLSSQVGLTSVTSIQ--ERIMST 467
QY 432 MSKAEIVQNLQNSKLYRLDNTWEEKLAKTEBIHKEREALAEELGTSIRK--GFVGPYH 489
Db 468 PGCEAIERLKESEKIITAEINLNETWEEKLRKTEAIRMERALLAEMGAIVREDGGTLGVFS 527
QY 490 SKEMPHLVNLSDDPLAECLVNIKPGOTRVGNVNDQAEIRLNGSKILKEHCTFENV- 548
Db 528 PKKTPHLVNLNEDPLMSECLLYIKGTRVQGAADARRQDIVLSGAHKEEHCFRSE 587
QY 549 ---DNVVTIVPNEKAAMVNVGRIDKPTRLSRGYRIILGDPHIFRHNHPEARAEQ 604
Db 588 SNSGEVITVLEPCERSETYVNGKRVSPQVQLRSGNRIIMGNHVFERNHPEQARAREK- 646
QY 605 SLLRHSVTNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRCKDSDFYARREAS 664
Db 647 -----TPSAETPSEPVDWTFAQRELE 668
QY 665 AILGLDQK-----ISHLTDDDELALFDD-----YOKARAVRGLVED 701
Db 669 K-QGIDMKQEMKRLQEMELIYKKEEADLLLEQORLDYESKLQALQKQVETRSLAAET 727
QY 702 NEDSDSOSFP 712
Db 728 TEEEBEEVEP 738

RESULT 15
US-60-130-217-6
; Sequence 6, Application US/60130217
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Dufaire-Gare, Isabelle
; APPLICANT: Greil, Pascale
; TITLE OF INVENTION: DNA encoding a novel kinesin-like protein
; FILE REFERENCE: GENSET.060PR
; CURRENT APPLICATION NUMBER: US/60/130,217
; CURRENT FILING DATE: 1999-04-20
```

```
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent.pm
; SEQ ID NO 6
; LENGTH: 1816
; TYPE: PRT
; ORGANISM: homo sapiens
us-60-130-217-6

Query Match      41.4%; Score 1668.5; DB 23; Length 1816;
Best Local Similarity 45.0%; Pred. No. 1.2e-121;
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;

Qy 4 GGNKVVVRVRFENAREIDRGAKICIVRMGNQITLTTPPGAEEKARKSGKTIIDGPKAFA 63
Db 3 GASVAVVRVRFENAREIDRGAKICIVRMGNQITLTTPPGAEEKARKSGKTIIDGPKAFA 63

Qy 64 FDRSYWSF--DKNAPNARYAQEDLFQDLGYPLLDNAFKYNNCIFYAGQTGSGKSYMMGYG 122
Db 52 FDYSWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNVCIFYAGQTGACKSYTMGCKQ 111

Qy 123 KEH--GVIPRICQDMFRINELQDKDKNITCTVEVSYLEIYNERVRDLLNPSTKGNLKYRE 180
Db 112 EESQAGIIPQCEELFEKIND--NCNEEMSYSEVSMEIYECYRVRDLNPKNKNLRYRE 170

Qy 181 HPSTGPVVEDLAKLVVRSFOEINLMDENKARTVAATNMNETSSRSRAVETLTITQKWH 240
Db 171 HPLGYPVEDLSKLAVSYTDIADLMDAGNKARTVAATNMNETSSRSRAVETLTITQKWH 230

Qy 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAENRSLSITLGRVIAALADM-- 298
Db 231 DNETNLSTEKYSKISLVLAGSERADSTGATGARLKEGANINKSLTTLGKVISALAEVDN 290

Qy 299 ---SSGKOKKQOLVYRSVLTWLLKDSLGNSMTAMTAISPADINFEETLSTLRYS 355
Db 291 CTSKSKKKKKTDFIPYRDSVLTWLLRENLGNSRTAMVAALSPADINDETLSLRYADR 350

Qy 356 AKRIKHAHVNEEDNARMIRELKEELAOLRSKLOSSGGG-----GGAG----- 399
Db 351 AKQICNAVINEEDNARMIRELKEEVLKOLLRAQGLGDIIDPLIDDDYGGSGSKYLK 410

Qy 400 -----GSGGPVEESYPPDTPLEKQ-----IVSIQQPDATVKK 431
Db 411 DFONKHRYLLASENORPGHFSTASMGSLTSS--PSSCSLSSQVGLTSVTSIQ--ERIMST 467

Qy 432 MSKAEIVQLNQSKLYRDLNOTWEEKLAKTEEIHKEKEAALEELGISIEK--GFVGYPH 489
Db 468 PGGEAIEERLKESEKIITAEELNWEELKRTATRMEREALLAEMGVAIREDGGTGLGVFS 527

Qy 490 SKEMPHLVNLSDDPLLAELCLVYNIKPGOTRYGVNQDTPQAEIRLNGSKILKEHCTFENV- 548
Db 528 PKTPHLVNLEDPLMSCELLYYIKDITRYGQADAEARRQDIVLSGAHIKBEHCIFRSER 587

Qy 549 ----DNVVTIVPNEKAAVMVNGVRIDKPTRLRSYRIILGDFHIFRNFHPEEAAERQEQ 604
Db 588 SNSGEVITLPECERSETYVNGKRVSPQLRSGNRIIMGNKHVFRFNHPQARAAREK- 646

Qy 605 SLLRHSVTNSQLSPAPGRHRTLSKAGSDADGDSRSPLPHFRGKDSDFYARREAAAS 664
Db 647 -----TPSAETPSEPVDWTFQAQRELE 668

Qy 665 AILGLDOK-----ISHLTDDDELDAFD-----VQKARAVRGLVED 701
Db 669 K-QGIDMKQEMEKRLQEMEILYKKEEADLLLEQQRLDYESKLOALQKQVETRSLAAET 727

Qy 702 NEDSDSSSFF 712
Db 728 TEEEEEVEP 738
```

THIS PAGE BLANK (USPTO)


```
Db 237 SVFSVTI-----HMKETTIDGEBLVKIGKLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292
Qy 286 STLGRVIAALADMSSGKQKKNQVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEE 345
Db 293 LTLGRVITALVE-----RTPHPYRESKLTIRIQLDSGGTRTSIIATISPASLNLEE 345
Qy 346 TSLTRYADSARIKRKNHVVNEDPNAR--MIRELKEELAQLRSKLOSGGGGAGGSGGP 404
Db 346 TSLTEYAHRAKNILNKPENVQKTKKALIKETYTEIERLKRDLAAAREKNG-----397
Qy 405 V---EESY-----PPDTPLEKQIVSQOPDATVKKMSKAEIVEQLNOSKLYRD----LNQ 453
Db 398 VYISEENFRVMSGKLTVOEQIVEL-----IEKIGAVE--EELNRVTLEFMDNKNELDQ 449
Qy 454 TWEELAKTEETH--KEREAALEELGISIEKGFVGPYHSKE 492
Db 450 CKSDLQNTQOLETTQKHLGETKGLVKEEYITSALESTE 489

RESULT 2
US-09-724-519-8
; Sequence 8, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-8

Query Match 16.7%; Score 673.5; DB 5; Length 575;
Best Local Similarity 36.0%; Pred. No. 3.4e-39;
Matches 187; Conservative 85; Mismatches 171; Indels 77; Gaps 18;

Qy 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTTMDGPKAPA 63
Db 16 GKNIQVVVRCRPFNLAERKASAHSIVECD-----PVRKEVSVRTGGLADKSKRKYT 67
Qy 64 FDRSWSGDKNAPNARQEDLFQDLGVPLLDNAPFGYNNCIPAYGOTSGKSYSMWG--- 120
Db 68 FDMVFGASTK-----QIDVRSVVCPLDEVIMGYNCTIFAYGOTGTGKFTMEGERS 120
Qy 121 ----YKKEH---GVIPRICQDMFRINELQKDKNLCTCTVEVSYLEIYNERYVRLINPST 172
Db 121 PNEEYTWEEEDPLAGIIPRTLHQIFEKLT----DNGETFSVKVLSLEIYNEELFDLLNPS 176
Qy 173 --KGNLKVREHP--STGYPVEDLAKLVRSFOEINLMDGKNKARTVAATNNNTESSRSH 228
Db 177 DVSERLQMDPDPNRKRGVILKGLEEITVHNKDEVYQILEKGAARKRTTAATLNAYSSRSH 236
Qy 229 AVFTLTQKHDEETKMDTE---KVAKISLDLAGSERATSTGATGARLKEGAENRSL 285
Db 237 SVFSVTI-----HMKETTIDGEBLVKIGKLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292
Qy 286 STLGRVIAALADMSSGKQKKNQVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEE 345
Db 293 LTLGRVITALVE-----HMKETTIDGEBLVKIGKLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292
Qy 286 STLGRVIAALADMSSGKQKKNQVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEE 345
```

```
Db 293 LTLGRVITALVE-----RTPHPYRESKLTIRIQLDSGGTRTSIIATISPASLNLEE 345
Qy 346 TSLTRYADSARIKRKNHVVNEDPNAR--MIRELKEELAQLRSKLOSGGGGAGGSGGP 404
Db 346 TSLTEYAHRAKNILNKPENVQKTKKALIKETYTEIERLKRDLAAAREKNG-----397
Qy 405 V---EESY-----PPDTPLEKQIVSQOPDATVKKMSKAEIVEQLNOSKLYRD----LNQ 453
Db 398 VYISEENFRVMSGKLTVOEQIVEL-----IEKIGAVE--EELNRVTLEFMDNKNELDQ 449
Qy 454 TWEELAKTEETH--KEREAALEELGISIEKGFVGPYHSKE 492
Db 450 CKSDLQNTQOLETTQKHLGETKGLVKEEYITSALESTE 489

RESULT 3
US-09-724-519-2
; Sequence 2, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-2

Query Match 16.3%; Score 658; DB 5; Length 1053;
Best Local Similarity 36.3%; Pred. No. 9.2e-38;
Matches 185; Conservative 83; Mismatches 163; Indels 78; Gaps 19;

Qy 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTTMDGPKAPA 63
Db 16 GKNIQVVVRCRPFNLAERKASAHSIVECD-----PVRKEVSVRTGGLADKSKRKYT 67
Qy 64 FDRSWSGDKNAPNARQEDLFQDLGVPLLDNAPFGYNNCIPAYGOTSGKSYSMWG--- 120
Db 68 FDMVFGASTK-----QIDVRSVVCPLDEVIMGYNCTIFAYGOTGTGKFTMEGERS 120
Qy 121 ----YKKEH---GVIPRICQDMFRINELQKDKNLCTCTVEVSYLEIYNERYVRLINPST 172
Db 121 PNEEYTWEEEDPLAGIIPRTLHQIFEKLT----DNGETFSVKVLSLEIYNEELFDLLNPS 176
Qy 173 --KGNLKVREHP--STGYPVEDLAKLVRSFOEINLMDGKNKARTVAATNNNTESSRSH 228
Db 177 DVSERLQMDPDPNRKRGVILKGLEEITVHNKDEVYQILEKGAARKRTTAATLNAYSSRSH 236
Qy 229 AVFTLTQKHDEETKMDTE---KVAKISLDLAGSERATSTGATGARLKEGAENRSL 285
Db 237 SVFSVTI-----HMKETTIDGEBLVKIGKLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292
Qy 286 STLGRVIAALADMSSGKQKKNQVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEE 345
Db 293 LTLGRVITALVE-----RTPHPYRESKLTIRIQLDSGGTRTSIIATISPASLNLEE 344
Qy 346 TSLTRYADSARIKRKNHVVNEDPNAR--MIRELKEELAQLRSKLOSGGGGAGGSGGP 404
```

```
Db 345 TLSTLEYAHRANKILNKPVNQKLTAKALIKETVEIERLKRDLAAREKNG----- 396
Qy 405 V---EESY----PPDTPLEKQVSIQOPDATVKKMSKAEIVQOLNQEKLRYD---LNQ 453
Db 397 VYISENFRVNSKLTVQEEQVEL-----IEKIGAVE--EELNRVTALFMDKNKELDQ 448
Qy 454 TWBEKLAKTEIHKERAALEELGISIEK 482
Db 449 CKSDLQNKTOEL-ETQKHLQETKLQLVK 476

RESULT 4
US-09-724-519-10
; Sequence 10, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Human
; US-09-724-519-10

Query Match 15.6%; Score 628; DB 5; Length 368;
Best Local Similarity 40.6%; Pred. No. 2.3e-36;
Matches 155; Conservative 57; Mismatches 122; Indels 48; Gaps 10;

Qy 4 GGNKVVVVRPFPNAREIDRGAKCIVRMEGNQITLTPPPGAEKARKSGKTMIDGPKAFA 63
Db 16 GKNIQVVRCPFNLAERKASAHSIVECD-----PVRKEVSVRTGGLADKSSRKYT 67
Qy 64 FDRSYWFDKNAPYARQEDLFODLGVPLLDNAFKGYNCIFAYGQTGSGKSYSMG--- 120
Db 68 FDMVFGASTK-----QIDVRSVVCPLIDEVINGYNTIFAYGQTGKTFTMEGERS 120
Qy 121 ----YKEH----GVIPRICODMFRINELQDKNLCTCTVEVSYLEIYNERVRLNPS 172
Db 121 PNEYTWEEEDPLAGIIPRTLHQIFEKL-----DNGTEFSVKVSLLEYNEELFDLLN 176
Qy 173 --KGNLKVREHP--STGPVVEDLAKLVRSFOEINLMDGKARTVAATNNNETSSRS 228
Db 177 DVSERLQMFDPNKRNGVYIKGLEEITVHNKDEYVIGLEGAARKTTAATLMNAYSRS 236
Qy 229 AVFTLTITQKHDEETKMDTE---KVAKISLDVLAGSERATSTGATGARLKEGAEINRS 285
Db 237 SVFESVTI----HMKETTIDGELVKIGKLNVDLAGSENIGRSGAVDKRAREAGNIQSL 292
Qy 286 STLGRVTAALADMSGKQKKNQVPRDSVLTWLLKDSLGNSMTAMIAAISPADINFE 345
Db 293 LTGLRVITALVE-----RTPHPVRESKLTIRLODSLGRTRTSIATISPASLNLE 345
Qy 346 TLSTLRVADSARKIKNHAVYNE 367
Db 346 TLSTLEYAHRANKILNKPVNQ 367

RESULT 5
```

```
US-09-724-519-4
; Sequence 4, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Human
; US-09-724-519-4

Query Match 15.3%; Score 618.5; DB 5; Length 382;
Best Local Similarity 39.4%; Pred. No. 1.1e-35;
Matches 157; Conservative 61; Mismatches 125; Indels 55; Gaps 12;

Qy 4 GGNKVVVVRPFPNAREIDRGAKCIVRMEGNQITLTPPGAEKARKSGKTMIDGPKAFA 63
Db 16 GKNIQVVRCPFNLAERKASAHSIVECD-----PVRKEVSVRTGGLADKSSRKYT 67
Qy 64 FDRSYWFDKNAPYARQEDLFODLGVPLLDNAFKGYNCIFAYGQTGSGKSYSMG--- 120
Db 68 FDMVFGASTK-----QIDVRSVVCPLIDEVINGYNTIFAYGQTGKTFTMEGERS 120
Qy 121 ----YKEH----GVIPRICODMFRINELQDKNLCTCTVEVSYLEIYNERVRLNPS 172
Db 121 PNEYTWEEEDPLAGIIPRTLHQIFEKL-----DNGTEFSVKVSLLEYNEELFDLLN 176
Qy 173 --KGNLKVREHP--STGPVVEDLAKLVRSFOEINLMDGKARTVAATNNNETSSRS 228
Db 177 DVSERLQMFDPNKRNGVYIKGLEEITVHNKDEYVIGLEGAARKTTAATLMNAYSRS 236
Qy 229 AVFTLTITQKHDEETKMDTE---KVAKISLDVLAGSERATSTGATGARLKEGAEINRS 285
Db 237 SVFESVTI----HMKETTIDGELVKIGKLNVDLAGSENIGRSGAVDKRAREAGNIQSL 292
Qy 286 STLGRVTAALADMSGKQKKNQVPRDSVLTWLLKDSLGNSMTAMIAAISPADINFE 345
Db 293 LTGLRVITALVE-----RTPHPVRESKLTIRLODSLGRTRTSIATISPASLNLE 345
Qy 346 TLSTLRVADSARKIKNHAVYNE 367
Db 346 TLSTLEYAHRANKILNKPVNQ 367

RESULT 6
US-09-723-153-2
; Sequence 2, Application US/09723153
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1048
; CURRENT APPLICATION NUMBER: US/09/723,153
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/634,957
; PRIOR FILING DATE: 2000-08-08
```

```
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Human
US-09-723-153-2

Query Match      12.7%; Score 513.5; DB 5; Length 790;
Best Local Similarity 26.2%; Pred. No. 4.2e-28;
Matches 193; Conservative 118; Mismatches 291; Indels 135; Gaps 24;

Qy 69 WSFQKAPNVARQEDL-FQDLGVPLLDNAFKYNNCFAYGOTGSGKSYMMGYG---KE 124
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 WSKLDGVLDASQDLVYETAKDVVSQALDGYNGTICYGOTGAGKTYTMMGATENTYKH 112
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 125 HGVIPRICQDMFRRINELQDKNLTCTVEVSYLEIYNERVRDILLN-----PSTKGNLKV 178
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 RGLPRLAQVFRMEERPTH---AIVRVSYLEIYNESLFDLLSTLPYVGPSTV-PMTI 168
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 179 REHSTGPGYVEDLAKLVRSFOETENLMDGNKARTVAATNMNETSSSHAVFTLTQK 238
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 169 VENPQ-GVFIKGLSVHLTSQEDAFSLFEGETNRITIASHTMKNKSSRSHCHIFTIYL--E 225
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 239 WHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTGRVIAALADM 298
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 AHSRTLSEKITYTSKINLVLDLAGSERLKGSGEQLKEATYINKSLFLEQAIILGAD- 284
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 299 SSGKQKNQLPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSAKR 358
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 ----QKRDHIPFRCKLTHALKDLSGGNCNMVLVTNIYGEAAOLEETLSLRFASRMKL 339
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 359 IKNIAVNEDPNA-RMIRELKEELAQLRKLSQSGGGGGAGGGGVEE----- 407
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 VTTEPAINEKYDAERVMKNLEKELALLKQELAIHDSLTNRFTVYDPMDEIQIAEINSQV 399
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 408 -SYPPDTPLEKQIVSIOQPDATVKMKSAEIVEQLNOSEKLYRDLNOTWEEKLAKTEIHH 466
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 RRYLEGLTDEIDIIISLQ-----IKEVFNQTRVLSQGEQVESTLRKRYTLI 447
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 467 KERE-----AALKEELG-ISIEKGFVGYHSEKEMPHLVNLSDDPLLAECILVYNIKPGQTRVG 521
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 448 DRNDFAAISAIOKAGLVDDVCHLVGEPEGQNF----GLGVAP-----FSTKPKG---- 492
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 522 NVNQDTQAEIRLNGSKILKECHCTENVDNVVTVIPNKEAAVWVNGVRID----KPTRLRS 577
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 493 -----KAKSKTKFKEPLS-----SLARKEGASSPVNGKDLDDYVYSTKSTQLVP 534
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 578 GYRIILGDFHIFRHNHPEAREARQEQSLRHSVTNSQLGSPAPGRHDRTLKAGSDADG 637
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 535 SSKD-GDV-----KDLMSRDRETSSTPELSDSPKEELRPIRPTPPSKPVAFEEF 584
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 638 DSRSDSLPHRPGKDSWFFYARREAASAILGDKOKISHLTDELDLDDVQKARAVRRG 697
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 585 KNEQGSSEINRIFKFNKSTLNERRKAS-----ETQOHNAIKRE 623
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 698 LVENEDSDSSSPVRKYMSNGTIDNFSLDTAITMPTGTPRSDDDGDLFFGDKKSKQD 757
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 IDVTKEALNFOKS--LREK---QGYENKGLMII-----DEERFLILKLDKLG 668
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 758 ASNVDDVEELRQQAQME 774
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 669 QYRSEYQDLRLRAEIQ 685
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-09-723-153-4
; Sequence 4, Application US/09723153
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
```

NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 4880891CD1
PCT-US01-08117-95

Query Match 3.3%; Score 133.5; DB 1; Length 1312;
Best Local Similarity 18.4%; Pred. No. 0.066;
Matches 168; Conservative 113; Mismatches 253; Indels 379; Gaps 43;

QY 60 KAFAPDRSYWSPDKNAPNVARQEDLPDGLVPLLDNAFGYN-NCIFAYGQTSGSKSYSM 118
DB 342 KLFRLVHLKGLGFD-NIESGAVWQVYDOLGIPVLSA-AGYNVK-
QY 119 MGVGKE-----HGVIPRICQDMFRRINEIQ-KDKNLCTCTVEVSYLEIYNE 162
DB 389 YLYGFEYCRSANTFOMALPEKVYVVKQKCEC-ENVKEIKVKEENET---EIKEIKMEE 444
QY 163 R-----VRDLLNPFSTGNLKVREHPSTG-----PYVEDLAKLV-VRSFOETE 203
DB 445 RNLIPEEKPIEDEI--ERKENIK-----PSLGSKKNLLESITPDSQKEVNIKKPEDNE 498
QY 204 NLMD-----EGNKARTVAATNMNETSSRSHAVFTLTITQKWHDEFTK 245
DB 499 NLDDKDDDTTRVDSNLNIKVEAEKAKSGDET'NKEEDEDDEA-----EIEEIEEIEE 552
QY 246 MOTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSITLGRVITAAALADSSGKOKK 305
DB 553 EDEDD-----DONNEEEFECYPPGKVKQ-----VRYGR-----GKQK 586
QY 306 NQLVPPYRDS-----VL-----TWLLKDSLGGNSMTAMTAAISPADINFEET 346
DB 587 MYEASIKDSVEGGEKVLVLYHYCGWNVRYDEWIKADKI-----VRPADKNVPR- 634
QY 347 LSTLYADSAKRKNHVVNEDPNARMIRELKEELAQLRSKLOSSGGGGAGGGGPGVPE 406
DB 635 -----IKHRKKINKLKEKXD-----EKYSPKCKLRL-----SKPPFQ 671
QY 407 ESYPPDTPLEKQIVSIOQPD-ATVKKMSKAEIVQOLNQSEKILRDLNOTWEEKLAKTEI 465
DB 672 TNPSPWYKSLDLDADKNSDTHAKSIETISILNGLQASESAEDSEQE--DERGAQMDN 730
QY 466 HKEREAALDELIGSIEKGFVGPYHSEMPHLVNLSDPLLAECLVYNIKPGQTRVGNVNO 525
DB 731 NGKESKID-----HLTNNDLISKE-----752
QY 526 DTQAEIRLNGSKILKECTFENVNVNVTIVPNEKAAVWYNGVRIDK-PTLRLSGYRIILG 584
DB 753 -----EQNSSLLLE-----NKVHADLVISKPYSKSPERLRKDIEVLSE 791
QY 585 DFH-----IFRPNHPEARAEQBSLRLHSVTNSQLGSPAPGRHD-- 625
DB 792 DTDYEEDVTKRKQVKKDQTDKSSKQIKRGR-----RYCNTECLATGSPCKKEK 845
QY 626 -----RTLKAGSDADG-----638
DB 846 ANKESLWMSNSSSDEDEETAKMTPTKYNGLKEKRKSLRTGTFYSGFSEVAKR 905
QY 639 ----SRSDPLPHFRKDSDFYARREASAILGLDQKISHLTDDLDLFDVQKARAV 694
DB 906 IKLLANSDERLONSRAKD-----RKDWSSIOGOWPK-----KTLKELFSDSDTEAAA 953
QY 695 -----RRGLVEDN-----EDSDSQS-----SEPRDK-----TMS 719
DB 954 SPHPAPEGVAEESQTVAEESCSVSVELEKPPPVNVDSKPIEEKTVEVNDKRAEPPS 1013
QY 720 NCTIDNFDLPTAITMCTPRSDDDGDLFFGDKKSKQDASNVVDEELRQOQAOQ----- 773
DB 1014 SGS--NSVLNTPPTTPESP-----SSVTTEGSRQSSSVTVSEPL 1051
QY 774 ----EALAKTAKOE 783
DB 1052 APNQBEVRSINKSE 1064

RESULT 9
PCT-US01-10484-145
Sequence 145, Application PC/TUS0110484
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-048
CURRENT APPLICATION NUMBER: PCT/US01/10484
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/668,680
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/695,618
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 09/728,711
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 172
SOFTWARE: Custom
SEQ ID NO 145
LENGTH: 2816
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-10484-145

Query Match 3.3%; Score 133.5; DB 1; Length 2816;
Best Local Similarity 16.7%; Pred. No. 0.19;
Matches 173; Conservative 159; Mismatches 309; Indels 397; Gaps 43;

QY 15 PFNAREIDRGACIIVMSEGNQITLTPPGAEEKARKSGKTIIMDGPKAFAPDRSYWSPDK- 73
DB 1730 PNNSRPLTPGT-----VYGGPPP-----AGAPWYGGPPPPNPFIPITPMGV 1771
QY 74 --NAPNY-----AROEDFODLGVPLLDNAFKYNNCIFAYGQTSGSKSYMMGYGKE 124
DB 1772 HCNPYEHNNLENEVSLEIDINQHL-----KSKREERWNRASKR 1810
QY 125 HGVIPRICQDMFRRINE-LQDKNLCTCTVEVSYLEIY-NERVRLDNLNTPSKGNLK----- 177
DB 1811 QS--EKMEELHNTDDLLQEKSECEVEELHRTVQKROQKDFID----GNVESLWTE 1864
QY 178 -----VREHPSTGYPVEDLAKLVRSFOEINLMDGCKARTVAATNMNETSSR----- 226
DB 1865 LEIEKSLKHHEDIVDEIECIEKTLKRRSELK-----EADRLAAESELSCTEKVKCLLV 1920
QY 227 ---SHAVFTLTITQKHDE-----ETKMDTEKV---AKISLVDLAGSERATSTGATGARL 275
DB 1921 VWGVSSVWTKNAVEFTDAKSLQTESDAEELERRAQTAVNLVKADQQLRSLOADAKD 1980
QY 276 KEGAEINRS--LSTLGRVIAAL-ADMSSGKQKKNQOLVPYRDSVLTWLLKDSLGGNSMTAM 332
DB 1981 LEQHKIQEELKINKIVAAKSDQCLSKKKEKIT-----EELQKLQKD----- 2026
QY 333 TAAISPADINFEETLSTLYAD--SAKR-----IKNHAVVNEDPNARMIREL--KEEL 381
DB 2027 ---TEAMERNEOHHLQVUKSEVLLQAKRAELEKLSQVTSQQQEMAVLDRQGHKKEEL 2083
QY 382 -----AQLRSKLQSSGGGG-----GAGSGGGVVEESYPPD 412
DB 2084 HLLOGSVQAKADLQEARLCEVTEKCNHIREVKSLLLEELSFQKGLNVLQISERKTOL 2143
QY 413 T----PLEKQIVSIQ-----OPDATVKKKMSKAE 436
DB 2144 TLIKQETEKENEENLQVLRQMSKHKTELKNILDMQLQENHELOGLKLQHDQVSELEKTO 2203
QY 437 T-----VEQLNQSEK-----LYRDLNQ----- 453

```
Db 2204 VAVLEEKLELENQAOISQOQKEIEMQKOLLERDKREITERMTABSRALQSCVECLSEKE 2263
Qy 454 -----TWEEKLAKT-----BEIHKEREAALFELGISIEK----- 482
Db 2264 DLQEKCDIWEKKLAQTKRVLAABEENSWEOSNLEKLEINVKLQOQLDQLNROKLSLHN 2323
Qy 483 -----GFYGP-----YHSEMPHLVNLSDPPLAELCLVYNI 513
Db 2324 DISAMQOQLDVLDTGLALTTSRRDGFSSPHFLPHKDLHLHTTK-HQDVLLSE----- 2376
Qy 514 KPGQTRV-----GNVNDQQA-----BIRLNGSKILKEHCTFENVNVV 552
Db 2377 ---QTRLOKDISEANRPEDCQBEETKQOQLVQLNEIEENKLVQOEMMFORLOKE- 2432
Qy 553 TIVPNEKAADVNGVRI-----DKPTRLRSYRIILGDFHIFRNPHEEAAE 600
Db 2433 --RESEESKLETSKVLKEQHLEKLTQDKSLDQVLSKVLAABEERVTLQEEERWCE 2490
Qy 601 RQEQSLLRHSVTNSOLGSPAPGRHRTLSKAGSDAGDSRSDSPHPRGKDSWFYARR 660
Db 2491 SLEXTL---SQTQKQSE---RQOLVKESGELLALQKEADSMRADFSLLRNQFLTERK 2543
Qy 661 EAASAILGLDQKI-----SHLTDDDELALFDD--VQKARAVRRLGVEDNED-----S 705
Db 2544 KAEQVASKALKATQRSQLEKNLLEQKQENSCIQKEMATTIELVAODNHERARRLMKELN 2603
Qy 706 DSQSFPPVRYKYMNGTTDNFSLDTAITMPTGPRSDDDGDALFGDKKSKQDASNVDEE 765
Db 2604 QMOVETEKKOMAN-----QKQD 2621
Qy 766 LRQOQAQMEELAKTAKQE 783
Db 2622 LERRQMEISDAMRTLKSE 2639
RESULT 10
US-60-248-505-1112
; Sequence 1112, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1112
; LENGTH: 992
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-1112
Query Match 3.2%; Score 130; DB 6; Length 992;
Best Local Similarity 19.0%; Pred. No. 0.076;
Matches 149; Conservative 128; Mismatches 286; Indels 222; Gaps 35;
Qy 85 FQD-LGVPL--LDNAFGKYNCFAYGOTGSKSYSMGYGKEHGV-----IP-----R 130
Db 55 FQDSQGVPLGTRNCTCFGVDBHLSFTGKRLS--SSAKMSGIGNKRAAGEPGTSHPEKKA 113
Qy 131 ICQDMFRINELQ-KDKNLCTCTEVSYLEIYNERYVRLDNLNSTKGNLKVREH---PSTGP 186
Db 114 AVESDGTTVETIKLGVSSSTELDIRTLQTKNRKLAEMLDQKQATEDELEHIEKLERRQ 173
Qy 187 YVEDLAKLVRSF-QETENL-----MDEG-----NKARTVAATNNNETSSRSHA 229
Db 174 ATDDASLLIVRYWSQFDENTRIILKRYDLEQGLDGLTERKALVVPPEPDSNQOE-- 231
Qy 230 VFTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINSLSTIG 289
```

```
Db 232 -----RKDRERKWEQBPAPFSFLATLASS-----SSEEMESQLQERVESR----- 274
Qy 290 RVIAALADAMSSKOKKNOLVPRDSVLTWLLKDSLGGGSMTAMIAAISPADINFETLST 349
Db 275 RAVSOITVYDKLQKVELLS-----RKLNGGDNLIV-----BEAVOE 312
Qy 350 LRYADSARKRIKNHAVNEDPNARMIRELKEE-----LAQLRSKLSQSGGGGGAGSGGPV 405
Db 313 L-----NSFLAQENMRLOELTDLLQEKHRPMSQEFKLSK-----V 349
Qy 406 BESYPPDPPLKQIVISIQOPDATYKMSK-----AEIVEQLN-OSEKLYRDLNNTWEEK 458
Db 350 ETAESRVLSMIDDLQWDIDKIRKREQLNRHLAEVLERNVSKGYKYGAGSSLYGCT 409
Qy 459 LAKTEEIKHEREAALEELGISIEKFGVGPYHSEMPHLVNLSDPPLAELCLVYNIKPGQT 518
Db 410 ITINARKFEEMNALEEE-----NKE-----LAQNRL---CELEKLRQDFE 446
Qy 519 RVGNVNDQTAQAEIRLNGSKILKEHCTFENVNVVTVIPNE-----KAAVMVNGVR- 568
Db 447 EYTONKELKVELRSAYEQVVKETPEYRCMQSQFSVLNYESQLKAHLDEARTLLHGTG 506
Qy 569 -----IDKPTR-----LRSGYRIILGDFHIFRNPHEEAAERQBSL-- 606
Db 507 THQHQVELIERDEVSLHKKLRTEVIQLEDTLAQVRKEYEMLRIEPEQTLLAANEQAGPINR 566
Qy 607 -LRISVT-----NSQLGSPAPGRHRTLSKAGSDAGDSRSDSPHPRGKDSWFYARR 660
Db 567 EMRLHLSLQNHHLQKEVIL-RYKRLKREAQSDLN-----KTRL 605
Qy 661 EAASAILGLDQKISHLTD--ELDALFDDVQKARAVRRLGVED-NEDSDSQSFF--V 713
Db 606 RSGSALLQSQSTEDPKDEPAELKPDSELSQSASAKASQEDANEIKKAQESQEMKLL 665
Qy 714 RDKYMSNCTIDNFSLDTAITMPTGPRSDDDGDALFGDKKSKQDASNVDEELRQQAQM 773
Db 666 LDMYRS-----APKEQDKVOLMAAEKKSK-----AELEDLRQLRXL 703
Qy 774 BEALK 778
Db 704 EDKEK 708
RESULT 11
US-60-248-505-895
; Sequence 895, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 895
; LENGTH: 3078
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-895
Query Match 3.1%; Score 126.5; DB 6; Length 3078;
Best Local Similarity 19.2%; Pred. No. 0.66;
Matches 147; Conservative 115; Mismatches 253; Indels 253; Gaps 38;
Qy 157 LEIYNERYVRLDNLNSTKGNLKVREHPTGPPYVEDLAKLVRSFQEIENLMDGNKARTVA 216
Db 747 LAVMSDKVKELCAKAEKLTLL---SHPSDAQIQEMKEDLVSSWEHIR----- 790
Qy 217 ATNNNETSSRHAVFTLTITQKWH-----DBETKMDTEKVAKISL---VDLAGSE--- 263
```


Db 214 KAQQWEEAAKOOIGASLKLPELREDEARVAAALQR-----LQIARTQLDDEANRLLRR 266
Qy 165 RDLN--NPSTGNLKVREHPSTGPPYVEDLAKLVVRSFOEINLM-----DEGKA 212
Db 267 RDELARRLSQJGEDIVREE-----RLVADNAQILARLDEEAEELLILSDSGRHADREMA 322
Qy 213 RTVAATNNETSSRSHAVFTLTLTK-----WHDEETKMD--TEKVAK 253
Db 323 FFAAAVKIAE-----SEAVFTSTAGRAEAAAGROQLERAIRDLSDRKLRKQSEASAE 378
Qy 254 ISLVD--LAG-----SERATSTGATGARLKEG---AEINRSLSLGRVIAALA--DMSSGK 302
Db 379 IDTIDKLSGLPDAERREAVEAAEIAVEDALIVAEAEAAVAEARSAAALARGPLEYAK 438
Qy 303 QKKNOLVPRDVSILTWLKDLSGGSWMTAMIAISPADINFEETLSTLYADSARKIKNH 362
Db 439 NRLNAL-DTEARTITKMLATSAAAGSTTPVAEEMTVERGYEAALG-----483
Qy 363 AVVNEDPNARMIRELKEELAQLRSKLOSS-----GGGGGAGGSGPVEESYPPDTPLEK 417
Db 484 AALGDD-----LESPLDASAPAYWGNGNGADDPLOQAK-----PL-- 521
Qy 418 QIVSQQPDATVKMKAIEIVBQLNQSEKLYRDLNQ-----TWEEKLAKTE-- 463
Db 522 -LDYAAQAPDALRRALAQIGVVADVSEARRLLPSLKAGQRLVTRREGALFRWDGHIAADAP 580
Qy 464 ---ETHKEREAALELGISIEKGFVGPYHSEKMPHLVNLSDPLLAELVLVNIKPGQTRV 520
Db 581 GAAALRLSQKNRLAEIAELD-----EARSILEAEEDQLAAK--TEDIRSELRL 628
Qy 521 GNVND-----TQAEIRLNGSKLKEHCTFFENVNVVITVPEKAAMVN 565
Db 629 SEVRDRSLATQLAEAREALTSAE-RASGDLRLRDVVVSEAQNOIGAQI--DEIAVOEE 685
Qy 566 GVRI--DKP-----TRLRGYRIILGFHIFRHNHPEEAERAEQESLLRHSVTNSQL 616
Db 686 NARIEMEDAPDLSVLDRLRRESQLEVATDRGLL-----AEARAREGVS---REAFSRQR 737
Qy 617 GSPAGRHDRT--LSKAGSDADGSDSPLPHFRGKDSDFWARREASAILGLDKTSH 675
Db 738 RIQATGOERSTWASRAAASAD-----HIATLREREER---AREEIAELDIAPEE-- 783
Qy 676 LTDDLDALFDVQKARAVRGLVEDNEDSDSQSFPPYVRDKYMSNGTIDNFSLDTAITMP 735
Db 784 -FDEKRRLLNLQKTEADARRAAADRLAEANLQR--AADRVAA-----TALS-- 828
Qy 736 GTPRSDDDGDALFFGDGKKSKODASNVVDVEELRQOQAQMEALKTAQE 783
Db 829 -----ELAEAREKGRAEERLVSAREK 850

RESULT 14
US-08-945-567C-4
; Sequence 4, Application US/08945567C
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-745 MIS
; CURRENT APPLICATION NUMBER: US/08/945,567C
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29

; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-08-945-567C-4

Query Match 3.0%; Score 120; DB 4; Length 1833;
Best Local Similarity 19.4%; Pred. No. 0.88;
Matches 160; Conservative 116; Mismatches 295; Indels 254; Gaps 46;
Qy 5 GNIKVVVRPNAREIDRGAKCIVRMENQVILP--PPGAEEKAR--KSKGTMDGPK 60
Db 560 GNVKI-----TNTGINAGGKAI-----TGLSPTLPISDIADQSSRIELGNTIQ---- 601
Qy 61 AFAPDRSYWSPDKNAPNVARQEDLFQDLGVPLLDNAF--KGVNCCI-FAYGQTGSGKSYS 117
Db 602 -----DKDKSNAASIND-----LLNTGFNLKNNNPIDFV-----STYD 635
Qy 118 MNGYKKEHGVIPRICODMFRINELQDKNL-TCTVEVSYLEIYNERNVRDLNLPSTKGNL 176
Db 636 IYDFANGNATTATVTHDTANKTSKVVDVNVDDTTIHL-----673
Qy 177 KVREHPSTGPPYVEDLAKLVRSFOEINLMDGKNKARTVAATNMNETSSRSHAVFTL--- 233
Db 674 -----TG--FDDNKLGVKT-TKLNKTSANGN-----TATNFVNSDEDALYNAKDI 718
Qy 234 -----TLQKWH-----DEETKMDTEKVAKISLVDLAGSERATSTGATGAR-----LKE 277
Db 719 AENLTLAKEHTTKGTADTALQTFVVKVDENNADANALTVGOKNANNQVNLTLKG 778
Qy 278 GAEINRSLSLGRVIAALADMSGGKKNQLVPRYRDSVLTLKLDKSLGNSMTAMIAATS 337
Db 779 ENGLAIKTDKNGTVTFGINTTSLGKAGKSTL--NDGGLS--IKNPTGEQIQV----- 827
Qy 338 PADINFEETLSLRVADSAAK--RIKNHAVVNE--DPNARMIRELKEELAQLRSKLOSSGG 393
Db 828 -----GADGVKFAKVNNGVVGAGIDGTTTRITR--DEI-----858
Qy 394 GGGAGGSGGPVEESYPPDTPLEKQIVSTQQPDANVKKMSKAEIVEOLNOS---EKLVRD 450
Db 859 ---GTTGTNGSLDKSKP---HLSKD--GINAGKKITNIQSGEIAQNSHDVATGKTIY-D 909
Qy 451 LNQTWEEKLAKTEEIHKBREAALEELGISIEKG---FVGPYHSEKMPHLVNLSDPLIA 506
Db 910 LKTELENKISSTA---KTAQNSLHEFSVADEQGNFTVSNPYSSYDTS---KTSDVITFA 963
Qy 507 --ECLVYNIKPGQTRGVNQNOD---TQAEIRL--NGSKILKE---HCTFFENVNVVTI 554
Db 964 GENGITTKVNGKVVVRVG-IDQTKGLTTPKLTIVGNNGKGVIVDSQNGQNTITGLSNTLAN 1022
Qy 555 VPNEKAAVMV---NGVRIDKPTRLSRGVRIILGFHI-----FRENH 593
Db 1023 VTNDKGSVRTTEQGNIIKDEDKTRAASIVDVLSAGFNLOGNCEAVDFVSTYDTVNFADGN 1082
Qy 594 PEEAAREQEQLLRHSV-----TNSQLGSPAPGRHRTLSKAGSDADGSDSDSPLP 646
Db 1083 ATTAKVTYDDTSKTSKVVDVNVDDTTIEVRDKLGVKTTTLTSTGTGANKFALS- 1138
Qy 647 HFRGKDSWFYARREAAASAILGLDQIKSHLTDELDALFDVQKARAVRGLVEDNEDS- 705
Db 1139 -----QATGDALVKASDIVAH-----LNTLSGDIQTA-----KGASQANNSAG 1176
Qy 706 --DSQSSFPVRD-----KV--MSGTID--NFSLDTAITMPTGP 738
Db 1177 YVDAGNKVIYDSTDNKYQAKNODGTVDKTEKVAKDLVAAQATP 1221

RESULT 15
US-08-945-567C-3
; Sequence 3, Application US/08945567C

GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-745 MIS
; CURRENT APPLICATION NUMBER: US/08/945,567C
; CURRENT FILING DATE: 1998-03-19
; PRIOR FILING DATE: 1998-03-19
; PRIOR FILING DATE: 1995-05-01
; PRIOR FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1996-03-26
; PRIOR FILING DATE: 1996-04-29
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
us-08-945-567C-3

Query Match 3.0%; Score 120; DB 4; Length 1992;
Best Local Similarity 19.4%; Pred. No. 0.99;
Matches 160; Conservative 116; Mismatches 295; Indels 254; Gaps 46;

Qy 5 GNIKVVVRFPNAREIDRGACIVRMEGNQTILTP--PPGAEEKAR--KSGKTIWMDGPK 60
Db 719 GNWKI-----TNTGINAGGKAI-----TGLSPTLPSTADQSSRNIELGNTIQ---- 760

Qy 61 AFADRSYMSFDKNAPYAROEFLFDLGVPLLDNAF--KGYNCI--FAYGTGSGKSYS 117
Db 761 -----DKDKSNAASIND-----ILNTGFNLKNNNPIDFV-----STYD 794

Qy 118 MMGYCKEHGVIPIRCODMFRINELQDKNL--TCVVEVSYLEYINERVRDLLNPSTKCNL 176
Db 795 IVDFAANGNATTATVTHDTANKTSKVYDVNVDDTTIHL----- 832

Qy 177 KVREHPSTGPVVEDLAKLVVRSFOEIEINLMDGNKARTVAATNNNETSSRSHAVFTL--- 233
Db 833 -----TG--TDDNKKLGVKI--TKLNKTSANGN-----TATNFVNSSDEDALYNAKDI 877

Qy 234 -----TLTQKWH-----DEETKMDTEKVAKISLVDLAGSERATSTGATGAR-----LKE 277
Db 878 AENLNTLAKEIHTTKGTADTALQTFYKKYDENNDADANAITVQKNANNQVNTLTLLKG 937

Qy 278 GAETNRSLSLTGRVIAALADMSGKKQKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAAIS 337
Db 938 ENGLNKTDKNGVTFTGINTTSGLKAGKSTL---NDGGLS--IKNPTGSEQIQV----- 986

Qy 338 PADINFETLSTLYADSAK--RIKNHAYVNE--DPNARMIRELKEELAQLSKIQSSGG 393
Db 987 -----GADGVKFAKVNNGNVVGVAGIDGTTTRITR---DEI----- 1017

Qy 394 GGGGAGSGGPPDEESYPPDTPLEKQIVSIQOPDATVKRMSKAEIVEQLNOS---EKLYRD 450
Db 1018 ---GFTGTNGSLDKSP---HLSKD--GINAGGKKTNIQSGEIAQNSHDVATGGKIY-D 1068

Qy 451 LQOTWEKLAKEEIHKEEAALELGISIEKG-----FVGPYHSEKMPHLYNLSDDPLLA 506
Db 1069 LATELENKISSTA---KTAQNSLHEFSVADEQGNFTVSNPSSYDTS---KTSDVITFA 1122

Qy 507 --ECLVYNIKPGQTRVGNVOD---TQAEIRL---NGSKILKE-----HCTFENVDNVVTI 554
Db 1123 GENGITTKVKNKGVVRVG-IDOTKGLTTPKLTGVNNGNKGIVIDSQNGQNTITGLSNTLAN 1181

Qy 555 VPNEKAAVMV-----NGVRIDKPTRLSRGYRIILGDPHI-----FRENH 593

Db 1182 VTNDKGSVRTTEQGNIIKDEKTRAASIVDVLISAGFNLOQNGEAVDFVSTYDTVYNFADGN 1241
Qy 594 PEEARAEKQFQSLLRHVS-----TNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLP 646
Db 1242 ATTAKVTYDDTSTKTSKVYDVNVDDTTIEVKDKKLGKVTTLTSTGTGANKFALSN---- 1297
Qy 647 HFRGKSDWFIYARREASAILGLDQKISHLTDDDELDAFDVGVOKARAVRGLVEDNEDS- 705
Db 1298 -----QATGDALVKASDIVAH-----LNTLSGDIQTA-----KGASOANNSAG 1335
Qy 706 --DSQSSFPVRD-----KY---MSGNTID---NFSLDTAITMPTGP 738
Db 1336 YVDADGNKVIYDSTDNKYQAKNDGTVDKTEVARDKLVAQAQTP 1380

Search completed: April 25, 2001, 10:16:29
Job time: 300 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:21 ; Search time 128.03 Seconds

(without alignments)
448.528 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357

Perfect score: 1834

Sequence: 1 MSGGNTKVVVRVPPNARE.....PADINFEETLSTLYRADSAR 357

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*

2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*

3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*

5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*

22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*

23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1834	100.0	784	1	PCT-US99-01355-1
2	1834	100.0	784	16	US-09-235-416-1
3	1834	100.0	784	20	US-09-654-850-1
4	1528.5	81.3	533	18	US-09-417-507-7298
5	1152	62.8	421	18	US-09-488-725A-6751
6	1152	62.8	1793	23	US-60-258-275-444
7	1149	62.6	893	18	US-09-488-725A-3179
8	1123	61.2	348	21	US-09-718-563-4
9	1123	61.2	348	21	US-09-718-804-4
10	1123	61.2	348	21	US-09-718-842-4

11	1123	61.2	1153	21	US-09-718-563-2	Sequence 2, Appli
12	1123	61.2	1153	21	US-09-718-804-2	Sequence 2, Appli
13	1123	61.2	1153	21	US-09-718-842-2	Sequence 2, Appli
14	1123	61.2	1770	21	US-09-718-563-6	Sequence 6, Appli
15	1123	61.2	1770	21	US-09-718-804-6	Sequence 6, Appli
16	1123	61.2	1770	21	US-09-718-842-6	Sequence 6, Appli
17	1119	61.0	354	21	US-09-718-563-10	Sequence 10, Appl
18	1119	61.0	354	21	US-09-718-804-10	Sequence 10, Appl
19	1119	61.0	354	21	US-09-718-842-10	Sequence 10, Appl
20	1119	61.0	1816	21	US-09-718-563-8	Sequence 8, Appli
21	1119	61.0	1816	21	US-09-718-804-8	Sequence 8, Appli
22	1119	61.0	1816	21	US-09-718-842-8	Sequence 8, Appli
23	1119	61.0	1816	23	US-60-130-217-6	Sequence 6, Appli
24	1117	60.9	1103	18	US-09-467-946-1	Sequence 1, Appli
25	1106.5	60.3	1773	23	US-60-167-217-18615	Sequence 18615, A
26	1106.5	60.3	1773	23	US-60-173-464-15257	Sequence 15257, A
27	1106.5	60.3	1773	23	US-60-191-637-18572	Sequence 18572, A
28	1106.5	60.3	1773	23	US-60-191-681-14693	Sequence 14693, A
29	1063	58.0	352	19	US-09-580-828-4	Sequence 4, Appli
30	1063	58.0	1362	19	US-09-580-828-2	Sequence 2, Appli
31	1028	56.1	1438	23	US-60-161-932-1631	Sequence 1631, Ap
32	1028	56.1	1921	23	US-60-167-217-15727	Sequence 15727, A
33	1028	56.1	1921	23	US-60-173-464-12890	Sequence 12890, A
34	1028	56.1	1921	23	US-60-191-637-15722	Sequence 15722, A
35	1028	56.1	1921	23	US-60-191-681-12455	Sequence 12455, A
36	987	53.8	359	21	US-09-721-689-4	Sequence 4, Appli
37	987	53.8	359	21	US-09-721-832-4	Sequence 4, Appli
38	987	53.8	359	21	US-09-722-139-4	Sequence 4, Appli
39	931.5	51.9	1375	21	US-09-721-689-2	Sequence 2, Appli
40	931.5	51.9	1375	21	US-09-722-139-2	Sequence 2, Appli
41	931.5	51.9	1375	21	US-09-722-139-2	Sequence 2, Appli
42	913.5	49.8	503	1	PCT-US00-09066-116	Sequence 116, App
43	911.5	49.7	504	1	PCT-US00-09066-115	Sequence 115, App
44	873	47.6	1174	23	US-60-167-217-11913	Sequence 11913, A
45	873	47.6	1174	23	US-60-173-464-9672	Sequence 9672, Ap

ALIGNMENTS

RESULT 1

PCT-US99-01355-1

; Sequence 1, Application PC/TUS9901355

; GENERAL INFORMATION:

; APPLICANT: Sakowicz, Roman

; APPLICANT: Goldstein, Lawrence S. B.

; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor

; TITLE OF INVENTION: Protein

; FILE REFERENCE: 18557C-000710PC

; CURRENT APPLICATION NUMBER: PCT/US99/01355

; CURRENT FILING DATE: 1999-01-22

; EARLIER APPLICATION NUMBER: US 60/072,361

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1

; LENGTH: 784

; TYPE: PRT

; ORGANISM: Thermomyces lanuginosus

; FEATURE:

; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed

; OTHER INFORMATION: microtubule motor protein

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (1)..(357)

; OTHER INFORMATION: kinesin-like microtubule motor domain

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (358)..(442)

; OTHER INFORMATION: neck domain links motor domain to stalk domain

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (713)
; OTHER INFORMATION: polymorphic variant #1 Val -> Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (762)
; OTHER INFORMATION: polymorphic variant #2 Asp -> Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (774)
; OTHER INFORMATION: polymorphic variant #3 Glu -> Asp
PCT-US99-01355-1

Query Match 100.0%; Score 1834; DB 1; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTTMDGPK 60
Db 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTTMDGPK 60

Qy 61 AFAFDRSWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQTSGSKYSMMG 120
Db 61 AFAFDRSWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQTSGSKYSMMG 120

Qy 121 YGKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDLPSTKGNLKYRE 180
Db 121 YGKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDLPSTKGNLKYRE 180

Qy 181 HPSTGPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
Db 181 HPSTGPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240

Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300
Db 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300

Qy 301 GKQKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357
Db 301 GKQKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357

RESULT 2
US-09-235-416-1
; Sequence 1, Application US/09235416A
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; FILE REFERENCE: 18557C-000710US
; CURRENT APPLICATION NUMBER: US/09/235,416A
; CURRENT FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: WO PCT/US99/01355
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 60/072,361
; EARLIER FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:

; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-235-416-1

Query Match 100.0%; Score 1834; DB 16; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTTMDGPK 60
Db 1 MSGGGNIKVVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTTMDGPK 60

Qy 61 AFAFDRSWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQTSGSKYSMMG 120
Db 61 AFAFDRSWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQTSGSKYSMMG 120

Qy 121 YGKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDLPSTKGNLKYRE 180
Db 121 YGKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDLPSTKGNLKYRE 180

Qy 181 HPSTGPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
Db 181 HPSTGPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240

Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300
Db 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300

Qy 301 GKQKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357
Db 301 GKQKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357

RESULT 3
US-09-654-850-1
; Sequence 1, Application US/09654850
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; FILE REFERENCE: 18557C-000710US
; CURRENT APPLICATION NUMBER: US/09/654,850
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/235,416
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 60/072,361
; PRIOR FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:

```

; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-654-850-1

Query Match      100.0%; Score 1834; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGGNIKVVVRPFNAREIDRGAKCIVRMEGNQITLTPPPGAEKARKSGKTINDGPK 60
Db 1 MSGGGNIKVVVRPFNAREIDRGAKCIVRMEGNQITLTPPPGAEKARKSGKTINDGPK 60
Qy 61 AFADRSYWSFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG 120
Db 61 AFADRSYWSFDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG 120
Qy 121 YGKEHGVIPRICODMFRINELQDKNLTCTVEVSYLEYINERVRDLINPSTKGNLKVRE 180
Db 121 YGKEHGVIPRICODMFRINELQDKNLTCTVEVSYLEYINERVRDLINPSTKGNLKVRE 180
Qy 181 HPSTGPYVEDLAKLVRSFOETENLMDGKNAKRTVAATNNMETSRSRSHAVFTLTQKWH 240
Db 181 HPSTGPYVEDLAKLVRSFOETENLMDGKNAKRTVAATNNMETSRSRSHAVFTLTQKWH 240
Qy 241 DETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTLGRVIAALADMS 300
Db 241 DETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTLGRVIAALADMS 300
Qy 301 GKQKNQLVPPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSTLYADSAK 357
Db 301 GKQKNQLVPPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSTLYADSAK 357

RESULT 4
US-09-417-507-37298
; Sequence 37298, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PAT99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 37298
; LENGTH: 533
; TYPE: PRT
; ORGANISM: A.fumigatus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (515)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-417-507-37298

Query Match      83.3%; Score 1528.5; DB 18; Length 533;
Best Local Similarity 86.7%; Pred. No. 2.1e-143;
Matches 294; Conservative 27; Mismatches 17; Indels 1; Gaps 1;

Qy 20 EIDRGAKCIVRMEGNQITLTPPPGAEKARKS-GKTINDGPKAFADRSYWSFDKNAPNY 78
```

```

Db 8 EIDRGAKCIVRMEGNQITLTPPPGAEKSKAGKGAVGPKTFAFDRSYWSFDKNAPNY 67
Qy 79 ARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWYKGEHGVIPRICODMFR 138
Db 68 AGQDNLFADLGVPLLDNAFQYNNCIFAYGOTGSGKSYSMWYKGEHGVIPRICODMFR 127
Qy 139 INEQKQKNLTCTVEVSYLEYINERVRDLINPSTKGNLKVREHPSGTPYVEDLAKLVRS 198
Db 128 TAKQEDKNLNTCTVEVSYLEYINERVRDLINPSTKGNLKVREHPSGTPYVEDLAKLVRS 187
Qy 199 FOETENLMDGKNAKRTVAATNNMETSRSRSHAVFTLTQKWHDEETKMDTEKVAKISLV 258
Db 188 PEEIDHLMDEGKNAKRTVAATNNMETSRSRSHAVFTLTQKWHDEETKMDTEKVAKISLV 247
Qy 259 LAGSERATSTGATGARLKEGAEINRSLSLTLGRVIAALADMSGKQKNQLVPPYRDSVLTW 318
Db 248 LAGSERANSTGATGARLKEGAEINRSLSLTLGRVIAALADVASGKKKNASWVYRDSILTW 307
Qy 319 LKDSLGNSMTAMIAAISPADINFEETLSTLYADSAK 357
Db 308 LKDSLGNSMTAMIAAISPADINFEETLSTLYADSAK 346

RESULT 5
US-09-488-725A-6751
; Sequence 6751, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FIPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL-genes_b Versions 1.0
; SEQ ID NO 6751
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-6751

Query Match      62.8%; Score 1152; DB 18; Length 421;
Best Local Similarity 62.1%; Pred. No. 5.4e-106;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;

Qy 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNQITLTPPPGAEKARKSGKTINDGPKAFA 63
Db 9 GASVKVAVRVPNSREMSRDKCIQMSGSTTTVNPQPKET-----PKSFS 57
Qy 64 FDRSYWSFDKNAP---NYARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG 120
Db 58 FDSYWS---HTSPEDINTASQKYRIDGEEMLQHAPEGYNVCIFAYGOTGAGKSYTMWG 115
Qy 121 YGK---EHGVIPRICODMFRINELQDKNLTCTVEVSYLEYINERVRDLINPSTKGNLKV 178
Db 116 KQEKDQGIIPQLCEDLSRINDFTND-NKSYSVESYMEIYCYERVLDLNPKNKGLRV 174
```

QY 179 REHPSTGYPVEDLAKLVRSFOEINLMDGKARTVAATNNNETSSSHAVFTLTQK 238
DB 175 REHPLGYPVEDLSKLVTSYNDIOLMDGKARTVAATNNNETSSSHAVFTLTQK 234
QY 239 WHDEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSLTGLGRVIAALADM 298
DB 235 RHDAETNITTEKVKISLVLAGSERADSTGAKGTRLEKAGANINKSLTTLGKVISALAE 294
QY 299 SSG-----KQKNQLVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLTLYA 353
DB 295 DSGPNKNNKKKTDPIPYRDSVLTWLLRENLGGSRTAWAALSADINFEETLTLYA 354
QY 354 DSAK 357
DB 355 DRK 358
RESULT 6
US-60-258-275-444
; Sequence 444, Application US/60258275
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001026-PROV
; CURRENT APPLICATION NUMBER: US/60/258,275
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 717
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 1793
; TYPE: PRT
; ORGANISM: HUMAN
US-60-258-275-444

Query Match 62.8%; Score 1152; DB 23; Length 1793;
Best Local Similarity 62.1%; Pred. No. 5.6e-105;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;
QY 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIMDGPKAPA 63
DB 9 GASVKVAVVRPFNSREMSRDSKCIHQMSGSTTTIVNPKQPKET-----PKSFS 57
QY 64 FDRSYWSDKNAP---NYARQEDLFQDLGVPLLDNAFGYNNCIPAYGQTGSGKSYMMG 120
DB 58 FDYSYWS--HTSPEDINVASQKQYRDIGEEMLQHAEGYNVCIFAYGQTGAGKSYTMG 115
QY 121 YGK--EHGVIPRICODMPFRINELQKDNLCCTVEVSYLEIYNERVRLNPNSTKGNLKV 178
DB 116 KOEKDQOGIIPQLCEDLFSRINDTND--NMSYSVEVSYMEIYCERVRLNPNKNGNLRV 174
QY 179 REHPSTGYPVEDLAKLVRSFOEINLMDGKARTVAATNNNETSSSHAVFTLTQK 238
DB 175 REHPLGYPVEDLSKLVTSYNDIOLMDGKARTVAATNNNETSSSHAVFTLTQK 234
QY 239 WHDEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSLTGLGRVIAALADM 298
DB 235 RHDAETNITTEKVKISLVLAGSERADSTGAKGTRLEKAGANINKSLTTLGKVISALAE 294
QY 299 SSG-----KQKNQLVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLTLYA 353
DB 295 DSGPNKNNKKKTDPIPYRDSVLTWLLRENLGGSRTAWAALSADINFEETLTLYA 354
QY 354 DSAK 357
DB 355 DRK 358
RESULT 7
US-09-488-725A-3179

; Sequence 3179, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 3179
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-3179

Query Match 62.6%; Score 1149; DB 18; Length 893;
Best Local Similarity 61.8%; Pred. No. 3.6e-105;
Matches 225; Conservative 59; Mismatches 56; Indels 24; Gaps 6;
QY 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIMDGPKAPA 63
DB 3 GASVKVAVVRPFNSREMSRDSKCIHQMSGSTTTIVNPKQPKET-----PKSFS 51
QY 64 FDRSYWSDKNAP---NYARQEDLFQDLGVPLLDNAFGYNNCIPAYGQTGSGKSYMMG 120
DB 52 FDYSYWA--HTSPEDINVASQKQYRDIGEEMLQHAEGYNVCIFAYGQTGAGKSYTMG 109
QY 121 YGK--EHGVIPRICODMPFRINELQKDNLCCTVEVSYLEIYNERVRLNPNSTKGNLKV 178
DB 110 KOEKDQOGIIPQLCEDLFSRINDTND--NMSYSVEVSYMEIYCERVRLNPNKNGNLRV 168
QY 179 REHPSTGYPVEDLAKLVRSFOEINLMDGKARTVAATNNNETSSSHAVFTLTQK 238
DB 169 REHPLGYPVEDLSKLVTSYNDIOLMDGKARTVAATNNNETSSSHAVFTLTQK 228
QY 239 WHDEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSLTGLGRVIAALADM 298
DB 229 RHDAETNITTEKVKISLVLAGSERADSTGAKGTRLEKAGANINKSLTTLGKVISALAE 288
QY 299 SSG-----KQKNQLVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLTLYA 353
DB 289 DSGPNKNNKKKTDPIPYRDSVLTWLLRENLGGSRTAWAALSADINFEETLTLYA 348
QY 354 DSAK 357
DB 349 DRK 352
RESULT 8
US-09-718-563-4
; Sequence 4, Application US/09718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-4

```

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-4

```

```

QY 123 KEH--GVIPRICQDMFRINELQKDNLTCTVEVSYLEIYNERVROLLNPSTKGNLKVRE 180
DB 111 EESQAGIIPOLCEELFEKIND-NCNEEMSYSEVSMEIYCERVOLLNPKNKNLVRRE 169
QY 181 HPSTGPPYVEDLAKLVVRSFOEINLMDGKARTVAATNNNETSSRSHAVFTLTITOKWH 240
DB 170 HPLLGPYVEDLSKLVTSYTDIADMDAGNKARTVAATNNNETSSRSHAVFTIVFTOKKH 229
QY 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLTGLRVTAALADMS 300
DB 230 DNETNLSTKSVKISLVLAGSERADSTGATGARLKEGAEINRSLTGLRVTAALADMS 289
QY 301 GKQKKNLQVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADSAK 357
DB 290 -KKKKTDFIPYRDSVLTWLLRENLGGSRTAMVAALSPADINDETSLTSLRYADRAK 345

RESULT 10
US-09-718-842-4
; Sequence 4, Application US/09/718842
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

```

```

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

Query Match      61.2%; Score 1123; DB 21; Length 348;
Best Local Similarity 61.6%; Pred. No. 3.1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

QY 4 GGNKVVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
DB 2 GASVAVVRVPRNAREIDRGAKCIVRMGNOTILTPPPGAEEKARKSGKTMIDGPKAPA 63
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

```

; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

Query Match 61.2%; Score 1123; DB 21; Length 1153;
Best Local Similarity 61.6%; Pred. No. 2.2e-102;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

```
Qy 4 GGNIKVVVRPFNAREIDRGAKCIIVRMGNOTILTPPGAEKARKSGKTIMDGPKAPA 63
Db 3 GASVAVVRPFNSRETSSKSKCIIQMGNSTSIINPKNPK-----APKSFS 51

Qy 64 FDRSYWSP--DRNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYMMGYG 122
Db 52 FDYSWHSHTSPEDPCFASQNRVYNDIGKEMLLHAPEGYNVCIFAYGQTGAGKSYTMGKQ 111

Qy 123 KEH--GVIPRICQDMFRINELQDKNLTCTVEYSYLEIYNRVDLLNPSTKGNLKYRE 180
Db 112 EESQAGIIPQCEELFEKIND--NCNEEMSYSEVSEIYECYERVDLLNPKNGNLVRE 170

Qy 181 HPSTGPYVEDLAKLVVRSFQETENIMDEGNKARTVAATNMNETSSRSRSHAVFTLTITQKWH 240
Db 171 HPLGPPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSRSHAVFTIVTQKHH 230

Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLUKEGAENRSLSTLGRVITAAADMS 300
Db 231 DNETNLSTEKYSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEVSK 290

Qy 301 GKQKNQLVPRYDVSVLTLWLLKDSLGGNSKMTAMIAISPADINFETLTSLRYADSAK 357
Db 291 -KKKKTDFIPYRDSVLTWLLRENLGNSRTAMVAALSPADINYEFTLTSLRYADRAK 346
```

```
RESULT 12
US-09-718-804-2
; Sequence 2, Application US/09/18804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-2
```

Query Match 61.2%; Score 1123; DB 21; Length 1153;
Best Local Similarity 61.6%; Pred. No. 2.2e-102;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

```
Qy 4 GGNIKVVVRPFNAREIDRGAKCIIVRMGNOTILTPPGAEKARKSGKTIMDGPKAPA 63
Db 3 GASVAVVRPFNSRETSSKSKCIIQMGNSTSIINPKNPK-----APKSFS 51

Qy 64 FDRSYWSP--DRNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYMMGYG 122
Db 52 FDYSWHSHTSPEDPCFASQNRVYNDIGKEMLLHAPEGYNVCIFAYGQTGAGKSYTMGKQ 111
```

```
Qy 123 KEH--GVIPRICQDMFRINELQDKNLTCTVEYSYLEIYNRVDLLNPSTKGNLKYRE 180
Db 112 EESQAGIIPQCEELFEKIND--NCNEEMSYSEVSEIYECYERVDLLNPKNGNLVRE 170

Qy 181 HPSTGPYVEDLAKLVVRSFQETENIMDEGNKARTVAATNMNETSSRSRSHAVFTLTITQKWH 240
Db 171 HPLGPPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSRSHAVFTIVTQKHH 230

Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLUKEGAENRSLSTLGRVITAAADMS 300
Db 231 DNETNLSTEKYSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEVSK 290

Qy 301 GKQKNQLVPRYDVSVLTLWLLKDSLGGNSKMTAMIAISPADINFETLTSLRYADSAK 357
Db 291 -KKKKTDFIPYRDSVLTWLLRENLGNSRTAMVAALSPADINYEFTLTSLRYADRAK 346
```

```
RESULT 13
US-09-718-842-2
; Sequence 2, Application US/09/18842
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-2
```

Query Match 61.2%; Score 1123; DB 21; Length 1153;
Best Local Similarity 61.6%; Pred. No. 2.2e-102;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

```
Qy 4 GGNIKVVVRPFNAREIDRGAKCIIVRMGNOTILTPPGAEKARKSGKTIMDGPKAPA 63
Db 3 GASVAVVRPFNSRETSSKSKCIIQMGNSTSIINPKNPK-----APKSFS 51

Qy 64 FDRSYWSP--DRNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYMMGYG 122
Db 52 FDYSWHSHTSPEDPCFASQNRVYNDIGKEMLLHAPEGYNVCIFAYGQTGAGKSYTMGKQ 111

Qy 123 KEH--GVIPRICQDMFRINELQDKNLTCTVEYSYLEIYNRVDLLNPSTKGNLKYRE 180
Db 112 EESQAGIIPQCEELFEKIND--NCNEEMSYSEVSEIYECYERVDLLNPKNGNLVRE 170

Qy 181 HPSTGPYVEDLAKLVVRSFQETENIMDEGNKARTVAATNMNETSSRSRSHAVFTLTITQKWH 240
Db 171 HPLGPPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSRSHAVFTIVTQKHH 230

Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLUKEGAENRSLSTLGRVITAAADMS 300
Db 231 DNETNLSTEKYSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEVSK 290

Qy 301 GKQKNQLVPRYDVSVLTLWLLKDSLGGNSKMTAMIAISPADINFETLTSLRYADSAK 357
Db 291 -KKKKTDFIPYRDSVLTWLLRENLGNSRTAMVAALSPADINYEFTLTSLRYADRAK 346
```

```
RESULT 14
US-09-718-563-6
; Sequence 6, Application US/09/18563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
```


; TITLE OF INVENTION: Novel motor proteins and methods for
; their use

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-6

Query Match 61.2%; Score 1123; DB 21; Length 1770;
Best Local Similarity 61.6%; Pred. No. 4.3e-102;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKVVVRPNAREIDRGAKCIVRMGNQTLTPPPGAEKARKSGKTMGPKAPA 63
Db 3 GASVKVAVRVPNSRETSKSKCIQMOGNSTSIINPKNPKE-----APKSF 51
Qy 64 FDRSYWSEF-DKNAPNTARQEDLFDLGVPLLDNAFKGYNNCIFAYGOTGSGKSYSMG 122
Db 52 FDYSWASHTSPEPCFASQNRVYNDIGKEMLLHAFEGYNYVCIFAYGOTGAGKSYTMGKQ 111
Qy 123 KEH--GVIPRICQDMFRRIEIQKDNLTCTVEVSYLEIYNERVRLDLPSTKGNLKVRE 180
Db 112 EESQAGIIPOLCEELFEKIND-NCNEEMSYSEVSYMEIYCERVDRDLNPKNKNLVRVE 170
Qy 181 HPSTGPPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTOKWH 240
Db 171 HPLGPPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQKH 230
Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADWSS 300
Db 231 DNETLSTEKYSKISLVDLAGSERADSTGKTRLKEGANINKSLTTLGKVISALAEVSK 290
Qy 301 GKQKNQLVYPRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLRVYADSAK 357
Db 291 -KKKKTDFIPYKDSVLTWLLRENIGGNSRTAMVAALSPADINDETLSTLRVYADRAK 346

RESULT 15

US-09-718-804-6
; Sequence 6, Application US/09718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-6

Query Match 61.2%; Score 1123; DB 21; Length 1770;
Best Local Similarity 61.6%; Pred. No. 4.3e-102;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKVVVRPNAREIDRGAKCIVRMGNQTLTPPPGAEKARKSGKTMGPKAPA 63
Db 3 GASVKVAVRVPNSRETSKSKCIQMOGNSTSIINPKNPKE-----APKSF 51
Qy 64 FDRSYWSEF-DKNAPNTARQEDLFDLGVPLLDNAFKGYNNCIFAYGOTGSGKSYSMG 122

Db 52 FDYSWASHTSPEPCFASQNRVYNDIGKEMLLHAFEGYNYVCIFAYGOTGAGKSYTMGKQ 111
Qy 123 KEH--GVIPRICQDMFRRIEIQKDNLTCTVEVSYLEIYNERVRLDLPSTKGNLKVRE 180
Db 112 EESQAGIIPOLCEELFEKIND-NCNEEMSYSEVSYMEIYCERVDRDLNPKNKNLVRVE 170
Qy 181 HPSTGPPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTOKWH 240
Db 171 HPLGPPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQKH 230
Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADWSS 300
Db 231 DNETLSTEKYSKISLVDLAGSERADSTGKTRLKEGANINKSLTTLGKVISALAEVSK 290
Qy 301 GKQKNQLVYPRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLRVYADSAK 357
Db 291 -KKKKTDFIPYKDSVLTWLLRENIGGNSRTAMVAALSPADINDETLSTLRVYADRAK 346

Search completed: April 25, 2001, 10:16:26
Job time: 327 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:29 ; Search time 7.94 Seconds
(without alignments)
141.828 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357

Perfect score: 1834

Sequence: 1 MSGGNIKVVVRPNARE.....PADINFEETSLRYADSAK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17899 seqs, 3154390 residues

Total number of hits satisfying chosen parameters: 17899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.New.*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	612	33.4	368	5	US-09-724-519-10
2	612	33.4	382	5	US-09-724-519-4
3	612	33.4	513	5	US-09-724-519-6
4	612	33.4	575	5	US-09-724-519-8
5	600.5	22.7	1053	5	US-09-724-519-2
6	456.5	24.9	338	5	US-09-723-153-4
7	456.5	24.9	790	5	US-09-723-153-2
8	94	5.1	331	5	US-09-739-449-10490
9	92.5	5.0	1118	6	US-09-739-449-9635
10	89	4.9	619	5	US-09-739-449-10081
11	86.5	4.7	636	5	US-09-739-449-9049
12	85	4.6	1008	5	US-09-308-453-2
13	84.5	4.6	605	5	US-09-739-449-9838
14	84.5	4.6	1312	1	PCT-US01-08117-95
15	83	4.5	1731	5	US-09-739-449-8331
16	82	4.5	3078	6	US-09-739-449-8331
17	82	4.5	3078	6	US-06-248-505-895
18	81	4.4	401	5	US-09-739-449-11084
19	80	4.4	377	5	US-09-739-449-12873
20	80	4.4	461	5	US-09-739-449-10886
21	77.5	4.2	521	5	US-09-739-449-9749
22	77	4.2	926	5	US-09-421-124-187
23	77	4.2	926	5	US-09-421-124-187
24	76.5	4.2	277	5	US-09-755-456-7
25	75.5	4.1	317	5	US-09-739-449-11446
26	75.5	4.1	492	5	US-09-739-449-13322
27	75	4.1	414	5	US-09-739-449-10769

28	75	4.1	471	5	US-09-739-449-9003	Sequence 9003, Ap
29	75	4.1	569	5	US-09-739-449-9683	Sequence 9683, Ap
30	74.5	4.1	420	5	US-09-739-449-9559	Sequence 9559, Ap
31	74	4.0	627	5	US-09-739-449-12517	Sequence 12517, Ap
32	73.5	4.0	299	5	US-09-739-449-10739	Sequence 10739, A
33	73.5	4.0	516	6	US-60-248-505-1279	Sequence 1279, Ap
34	73.5	4.0	686	5	US-09-739-449-11271	Sequence 11271, A
35	73.5	4.0	1156	5	US-09-661-322A-28	Sequence 28, Appl
36	73.5	4.0	1589	5	US-09-421-124-189	Sequence 189, App
37	73.5	4.0	1589	5	US-09-421-124-189	Sequence 189, App
38	73.5	4.0	3064	6	US-60-248-505-1172	Sequence 1172, Ap
39	73.5	4.0	3159	6	US-60-248-823-85	Sequence 85, Appl
40	73	4.0	221	5	US-09-739-449-12622	Sequence 12622, A
41	73	4.0	2031	5	US-09-739-449-10349	Sequence 10349, A
42	72.5	4.0	431	5	US-09-739-449-10147	Sequence 10147, A
43	72.5	4.0	551	5	US-09-818-143-21	Sequence 21, Appl
44	72.5	4.0	750	5	US-09-739-449-11486	Sequence 11486, A
45	72.5	4.0	979	5	US-09-739-449-10546	Sequence 10546, A

ALIGNMENTS

RESULT 1
US-09-724-519-10
; Sequence 10, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 05/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-10

Query Match 33.4%; Score 612; DB 5; Length 368;
Best Local Similarity 40.6%; Pred. No. 1.7e+46;
Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;

QY 4 GGNIKVVVRPNAREIDRCAGCIVRMENGTILTTPPGAEEKARKSGKTTMDGPKAFA 63
Db 16 GKNQVVRCPFNLAERKASHSIVECD-----PVRKEVSVRTGLADKSKRTYTT 67

QY 64 FDRYSWFDKNAPNVARQEDLFQDLGVPLLDNAFKGYNNCFAYGOTGSGKSYSMMG--- 120
Db 68 FDMVFGASTK-----QIDVRSVVCPIIDVINGYNTCTIFAYGOTGKTFTMEGERS 120

QY 121 ----YGRH----GVIPRICODMFRINELQDKNLCTCTVEVSVLEIYNERVRLNPST 172
Db 121 PNEYTWEEDPLAGIIPRTLQIQIEKLT----DNGTEFSYKVSLLIYNEELFDLLNPSS 176

QY 173 --KGNLKVREHP--STGPGYVEDLAKLVRSFOEINLMDGKNKARTVAATNNMTSSRSR 228
Db 177 DYVSRLQMFDDPNRKNRGIIRGLEETIVHNKDEVYGIKEGAARKRTAATLNNAYSSRSR 236

QY 229 AVFTLTITQKHDEETKMDTE---KVAKISLVDLAGSERATSTGATGARLKEGAENKSL 285
Db 229 :|||:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

Db 237 SVFSVTI-----HMKETTIDGELVKIGKLNVLVLAGSENIGRSGAVDKRAREAGNINQSL 292
Qy 286 STLGRVIAALADMSGKQKNQLVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFE 345
Db 293 LTLGRVITLVE-----RTPHVPYRESKULRIQDSLGGRTRTSIIATISPASINLEE 345
Qy 346 TLSTLRYADSAK 357
Db 346 TLSTLEYAHRK 357

RESULT 2

US-09-724-519-4
; Sequence 4, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-4

Query Match 33.4%; Score 612; DB 5; Length 382;

Best Local Similarity 40.6%; Pred. No. 1.8e-46;

Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;

Qy 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTTMDGPKAPA 63
Db 16 GKNIQVVVRCRPFNLAERKASAHSIVECD-----PVRKEVSVRTGGADKSSRKTYYT 67
Qy 64 FDRSYWSPDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG--- 120
Db 68 FDMVFGASTK-----QIDVYRSVVCPLIDEVINGYNTIFAYGOTGTGKFTTMEGERS 120
Qy 121 ----YGKEH-----GVIPRICDMFRINELQDKNLCTCTVEVSYLEIYNRVRLINPST 172
Db 121 PNEETWEEDPLAGIIPRTLHOIFEKLT-----DNGTEFSVKVSLLEIYNNEELFDLLNPSS 176
Qy 173 --KGMLKVREHP--STGPYVEDLAKLVRSFOEINLMDGKNKARTVAATNNNTESSRSH 228
Db 177 DVSERLQMFDDPPNRKRGVILKGLLEITVHNKDEVYGILEKGAARKRTTAATLMNAYSSRSH 236
Qy 229 AVFTLTQKWHDEETKMDTE---KVAKISLVDLAGSERATSTGATGARLKEGAENRSL 285
Db 237 SVFSVTI-----HMKETTIDGELVKIGKLNVLVLAGSENIGRSGAVDKRAREAGNINQSL 292
Qy 286 STLGRVIAALADMSGKQKNQLVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFE 345
Db 293 LTLGRVITLVE-----RTPHVPYRESKULRIQDSLGGRTRTSIIATISPASINLEE 345
Qy 346 TLSTLRYADSAK 357
Db 346 TLSTLEYAHRK 357

RESULT 3

US-09-724-519-6
; Sequence 6, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/724,519
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/592,037
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Human
US-09-724-519-6

Query Match 33.4%; Score 612; DB 5; Length 513;

Best Local Similarity 40.6%; Pred. No. 2.7e-46;

Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;

Qy 4 GGNIKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTTMDGPKAPA 63
Db 16 GKNIQVVVRCRPFNLAERKASAHSIVECD-----PVRKEVSVRTGGADKSSRKTYYT 67
Qy 64 FDRSYWSPDKNAPYARQEDLFQDLGVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG--- 120
Db 68 FDMVFGASTK-----QIDVYRSVVCPLIDEVINGYNTIFAYGOTGTGKFTTMEGERS 120
Qy 121 ----YGKEH-----GVIPRICDMFRINELQDKNLCTCTVEVSYLEIYNRVRLINPST 172
Db 121 PNEETWEEDPLAGIIPRTLHOIFEKLT-----DNGTEFSVKVSLLEIYNNEELFDLLNPSS 176
Qy 173 --KGMLKVREHP--STGPYVEDLAKLVRSFOEINLMDGKNKARTVAATNNNTESSRSH 228
Db 177 DVSERLQMFDDPPNRKRGVILKGLLEITVHNKDEVYGILEKGAARKRTTAATLMNAYSSRSH 236
Qy 229 AVFTLTQKWHDEETKMDTE---KVAKISLVDLAGSERATSTGATGARLKEGAENRSL 285
Db 237 SVFSVTI-----HMKETTIDGELVKIGKLNVLVLAGSENIGRSGAVDKRAREAGNINQSL 292
Qy 286 STLGRVIAALADMSGKQKNQLVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFE 345
Db 293 LTLGRVITLVE-----RTPHVPYRESKULRIQDSLGGRTRTSIIATISPASINLEE 345
Qy 346 TLSTLRYADSAK 357
Db 346 TLSTLEYAHRK 357

RESULT 4

US-09-724-519-8
; Sequence 8, Application US/09724519
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Beraud, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; TITLE OF INVENTION: proliferation states
; FILE REFERENCE: 1014A

; CURRENT APPLICATION NUMBER: US/09/724,519
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 09/592,037
 ; PRIOR FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: 09/428,156
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 575
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-724-519-8

Query Match 33.4%; Score 612; DB 5; Length 575;
 Best Local Similarity 40.6%; Pred. No. 3.2e-46;
 Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;
 QY 4 GGNKVVVRFPNAREIDRGAKCIVRMGNGQTLTPPPGAEKARKSGKTMIDGPKAFA 63
 Db 16 GKNQVVRCPFPNLAERKASAHISIVCD-----PVRKEVSVRTGGGLADKSSRKYT 67
 QY 64 FDRSYWSDKNAPNARQEDLFDQGLVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG--- 120
 Db 68 FDMVFGASTK-----QIDVRSVVCPIIDDEVIMGYNCTIFAYGOTGKTFTMEGERS 120
 QY 121 -----YGEH-----GVIPRICQDMFRINELQKKNLTCTVSVSLEYINERVDLLNPST 172
 Db 121 PNEYTWEEDPLAGIIPRLHQIFPEKLT-----DNGTEFVSVKVSLLEYINEELFDLLNPSS 176
 QY 173 --KGNLKVREHP--STGPPYVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSH 228
 Db 177 DVSRLOQFDDPRNKRGVIIKGLLEITVHNKDEVQIILEKGAARKTTAATLMNAYSRSRSH 236
 QY 229 AVFTLTQKHDEETKMDTE---KVAKISLVLDLAGSERATSTGATGARLKEGAINESSL 285
 Db 237 SVFSTVI-----HMKETIDGELVKIGLNLDLAGSENIGRSGAVDKRAREAGNINQSL 292
 QY 286 STLGRVIAALADMSGKOKKQNLVPRDSVLTMLLKDSLGNSMTAMIAISPADINFEF 345
 Db 293 LTLGRVITLVE-----RTPHPVRESKLTIRL-DSLGGRTTSIATISPASLNEE 344
 QY 346 TLSTLYADSAK 357
 Db 346 TLSTLEYAHRAK 357

RESULT 5
 US-09-724-519-2
 ; Sequence 2, Application US/09724519
 ; GENERAL INFORMATION:
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Mak, John
 ; APPLICANT: Sakowicz, Roman
 ; TITLE OF INVENTION: Methods of screening for modulators of
 ; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
 ; TITLE OF INVENTION: proliferation states
 ; FILE REFERENCE: 1014A
 ; CURRENT APPLICATION NUMBER: US/09/724,519
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 09/592,037
 ; PRIOR FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: 09/428,156
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1053
 ; TYPE: PRT
 ; ORGANISM: Human

US-09-724-519-2

Query Match 32.7%; Score 600.5; DB 5; Length 1053;
 Best Local Similarity 40.6%; Pred. No. 7.6e-45;
 Matches 151; Conservative 55; Mismatches 117; Indels 49; Gaps 11;
 QY 4 GGNKVVVRFPNAREIDRGAKCIVRMGNGQTLTPPPGAEKARKSGKTMIDGPKAFA 63
 Db 16 GKNQVVRCPFPNLAERKASAHISIVCD-----PVRKEVSVRTGGGLADKSSRKYT 67
 QY 64 FDRSYWSDKNAPNARQEDLFDQGLVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG--- 120
 Db 68 FDMVFGASTK-----QIDVRSVVCPIIDDEVIMGYNCTIFAYGOTGKTFTMEGERS 120
 QY 121 -----YGEH-----GVIPRICQDMFRINELQKKNLTCTVSVSLEYINERVDLLNPST 172
 Db 121 PNEYTWEEDPLAGIIPRLHQIFPEKLT-----DNGTEFVSVKVSLLEYINEELFDLLNPSS 176
 QY 173 --KGNLKVREHP--STGPPYVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSH 228
 Db 177 DVSRLOQFDDPRNKRGVIIKGLLEITVHNKDEVQIILEKGAARKTTAATLMNAYSRSRSH 236
 QY 229 AVFTLTQKHDEETKMDTE---KVAKISLVLDLAGSERATSTGATGARLKEGAINESSL 285
 Db 237 SVFSTVI-----HMKETIDGELVKIGLNLDLAGSENIGRSGAVDKRAREAGNINQSL 292
 QY 286 STLGRVIAALADMSGKOKKQNLVPRDSVLTMLLKDSLGNSMTAMIAISPADINFEF 345
 Db 293 LTLGRVITLVE-----RTPHPVRESKLTIRL-DSLGGRTTSIATISPASLNEE 344
 QY 346 TLSTLYADSAK 357
 Db 346 TLSTLEYAHRAK 357

RESULT 6
 US-09-723-153-4
 ; Sequence 4, Application US/09723153
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1048
 ; CURRENT APPLICATION NUMBER: US/09/723,153
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 09/634,957
 ; PRIOR FILING DATE: 2000-08-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 338
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-723-153-4

Query Match 24.9%; Score 456.5; DB 5; Length 338;
 Best Local Similarity 38.5%; Pred. No. 5e-33;
 Matches 115; Conservative 48; Mismatches 113; Indels 23; Gaps 8;
 QY 69 WSDKNAPNARQEDLFDQGLVPLLDNAFKYNNCIFAYGOTGSGKSYSMWG---KE 124
 Db 50 WSKLDGVLHDSODLVETVAKDQVSOALDGYNGTICMGOTGAGKTYTMGATENYKH 109
 QY 125 HGVIPRICQDMFRINELQKKNLTCTVSVSLEYINERVDLLN-----PSTKGNLKV 178
 Db 110 RGILPRALQVFRMEERPTH---AITRVSVSLEYINERVDLLN-----PSTKGNLKV 165
 QY 179 REHFTGPPYVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTQK 238
 Db 166 VENPQ-GVFIKGLSVHLTSQEDAFSLPPEGTNRIIASHTMKNKSSRSHCIFTIYL-E 222


```
QY 169 NPSTKGNLKVREH-----PSTGPPYVEDLA--KLVRVSFOEINLMDGCKARTV-AATNMNE 222
Db 221 -----RENGLOENNGQVLSLNGELNAQVLAQTOORLQOEESRLKQMEAAQNR 270
QY 223 TSSRSHAVFTLTOKWHEDETKMDT-----EKVAKISLVDLAGSERATSTGATGA 273
Db 271 T--QSDAIF-----DSOT-MNTIRAQYSTLIQQQIGAMTLTIGARHPLATAGAERA 318
QY 274 RLKEG-----AEINRSLTLGRVIAALADMSGGKKQKNQNLVPYR 313
Db 319 MLERGMADEARRILQAQKINVAERSSDALR--LKAVAEANVTDMDAQVRLRD 372

RESULT 10
US-60-248-505-785
; Sequence 785, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 785
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-785

Query Match 5.0%; Score 92.5; DB 6; Length 1118;
Best Local Similarity 22.1%; Pred. No. 1.2; Indels 65; Gaps 14;
Matches 57; Conservative 49; Mismatches 87;

QY 121 YGKEHGVPRICQ-----DMFRRNE-----LOKDKNLCT-VEVSYLEYINERVR 165
Db 821 FNKIHHELNAQCSAHTLQEVYIELFDIDENLKALQKQDLNMAPGLTIAQVVRTPKIP 880
QY 166 DLLNPSTKGNLKVREHSTGPPYVEDLAKLVRSFOE--IENLMDGCKARTVAAATNMNET 223
Db 881 EAI-----RNFELMEAKT-----KLLTAQKQKVKEKAEYTERKAVIEAKIAQ- 927
QY 224 SRSRSHAVFTLTOKWHEDETK---MDTEKVAKISLVDLA--GSRERATSTG-----A 270
Db 928 -----VAKIRPQKQVMEKETEKRISEIEGPRPLIVYVDAFLAREKAKADAEYVAHKA 981
QY 271 TGAR-----LKEGAEINRSLTLGR-VIAALADMSGGKKQKNQNLVPYRDSVLTWLL 320
Db 982 TSNKHLTPLEYELKKVQAINR-----LKGALSKLSGHTADSRKISPLEPGGAAPLIWT 1037
QY 321 KDSLGNSMTAMIAISP 338
Db 1038 KOVEG--SSNAMEAAVSP 1053

RESULT 11
US-09-739-449-10081
; Sequence 10081, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 10081
; LENGTH: 619
```

```
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10081

Query Match 4.9%; Score 89; DB 5; Length 619;
Best Local Similarity 22.3%; Pred. No. 1.1; Indels 84; Gaps 16;
Matches 75; Conservative 45; Mismatches 133;

QY 78 YARQ-EDLFDQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSY-----SMGCGY-KEHGVIPRI 131
Db 48 YQOVQSDLSDFQFTRLTTEQAKAFGIDRQVFSQLAAGAALDELASKMNLGLSENRLANLI 107
QY 132 CQD-MFRRIEQLQKNLTCTVEVSYLEYINERVDLLNPSTKGNLKVREHSTGPPYVED 190
Db 108 AEDPAKSVNG-QFDRNL-----FSERLR-----NSGFR-----DDYIKE 142
QY 191 LAKLVVRSFOEINLMDGCKARTVAAATNMNETSSRSHAVFTL-----T 234
Db 143 RSKVAIRS--QIVEAVSDGFAAPQVLDALKQYRNEQRAVDYVILSNVJPPVKAPGDDV 200
QY 235 LTQKWHDETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINR-SLSTLGR--- 290
Db 201 LTPPEWTNKSRYRAPEFRKFTYVKLEPDSIAEPASVTDQAIDYIYNHSHKDSFRTAGRRTV 260
QY 291 -----VIAALADMSGGKKQKNQNLVPYR-----DSVLTWLLKDSLGNSMTAMIAA 335
Db 261 EQLTPPDKEMAAAAAEQIRLNTTYDQVVKDQGTASDVTLGEFTKDTIPDQSIADAFA 320
QY 336 I-----SP-ADINFEETL-----STLRVADSAS 357
Db 321 IQKGGVSPVVDGSPGPIILLRVYTGKIPETTRTLDEAK 357

RESULT 12
US-09-739-449-9049
; Sequence 9049, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9049
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9049

Query Match 4.7%; Score 86.5; DB 5; Length 636;
Best Local Similarity 24.3%; Pred. No. 1.9; Indels 63; Gaps 18;
Matches 82; Conservative 40; Mismatches 153;

QY 17 NAREIDRCAGKACIVRMENGTILTPPPGAEEK--AR-KSGKTIKD-----GPKAFAP----- 64
Db 162 NAGALD--ARLIVLNDMDMSTAPPTGMSAYLARIASGRTYMGFRDFGKKLTAYLGKTI 219
QY 65 DRSYWSFDKAPNRYARQEDLFQDLG-----VPLLDNAFKGYNNCIFAYGQ 109
Db 220 DRAITRAVTHARGYVGTGTLPELGFVHIGIDHGSFDBLLPVLNRVRDQKGPVLIHW 279
QY 110 TCGSGKSY--SMGCGYKEHGVIPRICQMDPRINELQKKNLTCTVEVSYLEYINERVDL 167
Db 280 TQKGYAPAEAAADRYKHGV-----NKFVDITGQAQAKAPNAP---SYTSVFAGAE---L 327
QY 168 LNPSTKGN--LKVRHPSSTGPPYVEDLAKLV-VRSFOEINLMDG-----NKAETVAATNMN 221
Db 328 IQEATLDEKIIGVTAAMPNGTGLQKMAELFPSTF-----DVGIAEQHAVTFAA-GIA 379
```

```
QY 222 ETSRSRSHAVFTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEI 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 ADGYKPFICALYSTFLQGVYDQL-----VHDVAIQSLPVPRFPIDRAGEVGDGP-THAGSPD 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 282 NRSSTL-GRVTAALADMSSGKQKNQNLVYPYRDSVLTW 318
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 TTFLATLPGMVVYMAAADEAEELKHMYRTRAAAYDEGPISF 472
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-09-308-453-2
; Sequence 2, Application US/09308453
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GMBH
; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and
; TITLE OF INVENTION: for isolating cells and groups of cells
; FILE REFERENCE: BMD9924US
; CURRENT APPLICATION NUMBER: US/09/308,453
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Clostridium histolyticum
US-09-308-453-2

Query Match 4.6%; Score 85; DB 5; Length 1008;
Best Local Similarity 19.0%; Pred. No. 4.8;
Matches 63; Conservative 40; Mismatches 143; Indels 86; Gaps 13;

QY 53 KTIMDGPFAFADRSYWSF-----DKNAPNPTARQEDLFDQLGVPLLDNAFK-- 98
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 KTLNSG-----YDSDWMFYNTGFAVAHYLYEKDMPTEFKMKKAILNTDVKSYDEIKKL 537
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 -----GYNNCI--FAYGQTGSGKSYSMGYGKEHGVIPRICODMPRINELQKKNL 148
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 538 SDANKNTEYQHIEQLVDYKYGAGLPLVSDYLDKHGY--KKASEVYSEISKAASLTWT 595
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 149 TCTVEVS-YLEYINRVRDLNLPSTKGNLKVREHPSTGYPVEDLAKLVRSFQEIENLMD 207
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 596 SVTAEKSYQVNTFLR-----GTYTGTSKGEFKDWDENSKKLD 634
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 208 -----EGNKARTVAATNMNETSSRS---HAVFTLTQKWHDEETKMDTEKVAK 253
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 GTLESKXNSGKYTLTAFTNYRVTSDNKVQYDVVFHGVLTNDGDISNNKAPITAKVT- 693
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 ISLVDLAGSERATSGATGARLKEGAEINRSLSTLGRVIAALADMSSG--KOKKNOLVPY 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 694 -----GPSTGAVGRNIEFGSKDSKDED--GKIVSYDWDPDGDGATSRGKNSVHAY 740
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 RDS-----VLTWLKDSLGGNSMTAMIAISPAD 340
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 741 KKAGTYNVLTKVTDKDATATESFTEIKNED 772
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-739-449-9838
; Sequence 9838, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9838
; LENGTH: 605
```

```
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9838

Query Match 4.6%; Score 84.5; DB 5; Length 605;
Best Local Similarity 20.4%; Pred. No. 2.6;
Matches 82; Conservative 59; Mismatches 165; Indels 95; Gaps 19;

QY 5 GNKVVVVRPFENAREIDRGAKCIVRMEG-----NQTLTPPPGAEBEAKRSKKTMDGPK 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 244 GSMAAAVAVFRSNAME-----RLRLEGDAEQNRTL--SEQERNERERTAAKDAAD--I 292
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AFADRSYWSFDPKNAPNVARQEDLFDQLGVPL--LDNAFKGYNNCIFYAGQTGSGKSY 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 QFAYD-----SLAKGLAHL-DCGLNVRIDTPTFVTRIDRDNFNNSVAKLNALSTVGQN 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 MMGYGKEHGVIPRICODMPRINELQKD-----KNLTCTVEVSYLEIYNRVRDLN 169
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 ALAIDAGAGEIRQSADDLARRTEQQAASVEETAALAEITTVKDSARRA--EEVGRLYD 405
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 PSTKGNLKVREHPSTGYPVEDLAKL---VVRSFQEIENLM----- 206
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 406 -RARGNAE-----QSGVIVEDAVRAMEGIEKSSSEISNIIGVIDEIAFOTNLLNAGVE 459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 ----DEGNKARTVAATNMNETSSRS---HAVFTLTQKWHDEETKMDTEKV-AKISLV 257
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 460 AARAGEAGGFVAVQAEVRELAORSANAAKAIKTL-----INASTSQVQSGVELV 509
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 DIAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKOKKNOLVPRD--- 313
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 510 GNAGKALETI-----VREVEQEIHRHDA---IVTSREQSTGLOEINTAINTIDOGTQ 559
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 SVLTWLKDSLGGNSMTAMIAISPADINFEETLSTLYAD 354
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 560 QNAAMVEEOTAASHGLASEAAALNELLAQFLAAATRRQAE 600
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
PCT-US01-08117-95
; Sequence 95, Application PC/TUS0108117
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAL, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/08117
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 95
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 4880891CD1
PCT-US01-08117-95

Query Match 4.6%; Score 84.5; DB 1; Length 1312;
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:26 ; Search time 128.03 Seconds
(without alignments)
229.918 Million cell updates/sec

Title: US-09-235-416-1-copy_602_784

Perfect score: 938
Sequencing: 1 QPSQLLRHSVNSQLGSPAP.....ELRQQQAQMEBALTKAQEF 183

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%
Maximum Watch 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	100.0	784	1 PCT-US99-01355-1	Sequence 1, Appli
2	938	100.0	784	16 US-09-235-416-1	Sequence 1, Appli
3	938	100.0	784	20 US-09-654-850-1	Sequence 1, Appli
4	472.5	50.4	1000	18 US-09-417-507-42125	Sequence 42125, A
5	102	10.9	466	18 US-09-417-507-33143	Sequence 33143, A
6	99	10.6	845	21 US-09-733-089-17684	Sequence 17684, A
7	98.5	10.5	1030	7 US-08-302-166-4	Sequence 4, Appli
8	98.5	10.5	1735	23 US-60-173-464-5644	Sequence 5644, Ap
9	98.5	10.5	1736	23 US-60-167-217-7043	Sequence 7043, Ap
10	98.5	10.5	1883	23 US-60-191-637-6927	Sequence 6927, Ap

11	98.5	10.5	1883	23	US-60-191-681-5421	Sequence 5421, Ap
12	98	10.4	414	16	US-09-248-796-19046	Sequence 19046, A
13	97	10.3	775	21	US-09-733-089-17565	Sequence 17565, A
14	97	10.3	775	21	US-09-733-089-17566	Sequence 17566, A
15	94	10.0	143	23	US-60-197-873-25538	Sequence 25538, A
16	93.5	10.0	972	23	US-60-173-464-21969	Sequence 21969, A
17	93.5	10.0	972	23	US-60-191-637-31603	Sequence 31603, A
18	93.5	10.0	972	23	US-60-191-681-25084	Sequence 25084, A
19	93	9.9	842	16	US-09-248-796-15605	Sequence 15605, A
20	92.5	9.9	264	1	PCT-US00-05882-1095	Sequence 1095, Ap
21	92.5	9.9	676	18	US-09-450-969-4642	Sequence 4642, Ap
22	91	9.7	219	23	US-60-140-109-12	Sequence 12, Appli
23	91	9.7	842	23	US-60-259-128-4630	Sequence 4630, Ap
24	89.5	9.5	590	12	US-08-893-852-4	Sequence 4, Appli
25	89.5	9.5	590	14	US-09-052-753-2	Sequence 2, Appli
26	89.5	9.5	590	14	US-09-052-753-2	Sequence 2, Appli
27	89.5	9.5	1028	7	US-08-307-742-4	Sequence 4, Appli
28	89.5	9.5	1115	1	PCT-US94-10230-2	Sequence 2, Appli
29	89.5	9.5	1115	8	US-08-487-826-2	Sequence 2, Appli
30	89.5	9.5	1115	8	US-08-487-826-2	Sequence 2, Appli
31	89.5	9.5	1115	9	US-08-568-459-2	Sequence 2, Appli
32	89.5	9.5	1115	16	US-09-210-288-2	Sequence 2, Appli
33	89.5	9.5	1245	10	US-08-605-185-4	Sequence 4, Appli
34	89	9.5	420	21	US-09-733-089-13868	Sequence 13868, A
35	88	9.4	631	23	US-60-191-637-26902	Sequence 26902, A
36	88	9.4	631	23	US-60-191-681-21511	Sequence 21511, A
37	88	9.4	1530	23	US-60-240-409-736	Sequence 736, App
38	87.5	9.3	532	18	US-09-417-507-30854	Sequence 30854, A
39	87.5	9.3	933	14	US-09-095-272-2	Sequence 2, Appli
40	87.5	9.3	1021	23	US-60-269-308-4037	Sequence 4037, Ap
41	87	9.3	476	17	US-09-311-894-12	Sequence 12, Appli
42	87	9.3	936	11	US-08-781-986A-5249	Sequence 5249, Ap
43	87	9.3	936	13	US-08-956-171-5249	Sequence 5249, Ap
44	87	9.3	936	13	US-08-956-171B-5249	Sequence 5249, Ap
45	86.5	9.2	340	1	PCT-US97-21783-187	Sequence 187, App

ALIGNMENTS

RESULT 1
PCT-US99-01355-1
; Sequence 1, Application PC/TUS9901355
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; FILE OF INVENTION: Protein
; CURRENT APPLICATION NUMBER: PCT/US99/01355
; EARLIER FILING DATE: 1999-01-22
; EARLIER FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; FEATURE:
; NAME/KEY: DOMAIN

; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (713)
; OTHER INFORMATION: polymorphic variant #1 Val -> Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (762)
; OTHER INFORMATION: polymorphic variant #2 Asp -> Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (774)
; OTHER INFORMATION: polymorphic variant #3 Glu -> Asp
PCT-US99-01355-1

Query Match 100.0%; Score 938; DB 1; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEQSLRHVSNTSOLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFRGKDSDFYARRE 60
Db 602 QEQSLRHVSNTSOLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFRGKDSDFYARRE 661
|||
QY 61 AASAILGLDQKISHLTDELDFDDVQKARAVRGLVEDNEDSDSQSFPPVRDKYMSNG 120
Db 662 AASAILGLDQKISHLTDELDFDDVQKARAVRGLVEDNEDSDSQSFPPVRDKYMSNG 721
|||
QY 121 TIDNFSLDTAITMPGTPRSDDGDLFFGDKKSKQDASNVDELROOQQAQMEALKTAK 180
Db 722 TIDNFSLDTAITMPGTPRSDDGDLFFGDKKSKQDASNVDELROOQQAQMEALKTAK 781
|||
QY 181 QEF 183
Db 782 QEF 784
|||

RESULT 2
US-09-235-416-1
; Sequence 1, Application US/09235416A
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; FILE REFERENCE: 18557C-000710US
; CURRENT APPLICATION NUMBER: US/09/235,416A
; CURRENT FILING DATE: 1999-01-22
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: WO PCT/US99/01355
; EARLIER APPLICATION NUMBER: US 60/072,361
; EARLIER FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
; NAME/KEY: DOMAIN

; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-235-416-1

Query Match 100.0%; Score 938; DB 16; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEQSLRHVSNTSOLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFRGKDSDFYARRE 60
Db 602 QEQSLRHVSNTSOLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFRGKDSDFYARRE 661
|||
QY 61 AASAILGLDQKISHLTDELDFDDVQKARAVRGLVEDNEDSDSQSFPPVRDKYMSNG 120
Db 662 AASAILGLDQKISHLTDELDFDDVQKARAVRGLVEDNEDSDSQSFPPVRDKYMSNG 721
|||
QY 121 TIDNFSLDTAITMPGTPRSDDGDLFFGDKKSKQDASNVDELROOQQAQMEALKTAK 180
Db 722 TIDNFSLDTAITMPGTPRSDDGDLFFGDKKSKQDASNVDELROOQQAQMEALKTAK 781
|||
QY 181 QEF 183
Db 782 QEF 784
|||

RESULT 3
US-09-654-850-1
; Sequence 1, Application US/09654850
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; FILE REFERENCE: 18557C-000710US
; CURRENT APPLICATION NUMBER: US/09/654,850
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/235,416
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 60/072,361
; PRIOR FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-654-850-1

```

Query Match          100.0%; Score 938; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. NO. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEQSLRHVSYNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFHFGKDSDFYARRE 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 QEQSLRHVSYNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFHFGKDSDFYARRE 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 AASAILGLDQKISHLTDBDELDAFDDVQKARAVRGLVEDNEDSDSQSFVPVRDKYMSNG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 AASAILGLDQKISHLTDBDELDAFDDVQKARAVRGLVEDNEDSDSQSFVPVRDKYMSNG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 TIDNFSLDTAITMPTGPRSDDDGDALFPGDKSKQDASNDVDVELRQOQAOMEALKTAK 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 TIDNFSLDTAITMPTGPRSDDDGDALFPGDKSKQDASNDVDVELRQOQAOMEALKTAK 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 QEF 183
    |||
Db 782 QEF 784
    |||

RESULT 4
US-09-417-507-42125
; Sequence 42125; Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 42125
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: A.fumigatus
US-09-417-507-42125

Query Match          50.4%; Score 472.5; DB 18; Length 1000;
Best Local Similarity 57.4%; Pred. NO. 1e-38;
Matches 101; Conservative 22; Mismatches 28; Indels 25; Gaps 4;

QY 32 SDADGD--SRSDSPLPHFGKDSDFYARREASAILGLDQKISHLTDDDELDAFDDVQKA 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 SELDGDSSRADSPQAQRGGEADWFTARREAVSAILGPDH-ISHMPDDELDAFDDVQKV 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 91 RAVRGLVEDNEDSDSQSFVPVRDKYMSNGTTIDNFSLDTAITMPTGPRSDDDGDALFFGD 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 RATARGLVENEDSDSLSFVPVRDKYMSNGTTIDNFSLDTAITMPTGPRQQQYDGGGNGSD 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 ---RKSKQD-----ASNDVVEELRQOQAOMEALKTAKQEF 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 FTLQAARQDQRHLDKQEEFKNKLRIAEASDDADELRLEKEREALRSTKEEY 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-09-417-507-33143
; Sequence 33143; Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 33143
; LENGTH: 466
; TYPE: PRT
; ORGANISM: A.fumigatus
; FEATURE:

```

```

; NAME/KEY: UNSURE
; LOCATION: (412)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-09-417-507-33143

Query Match 10.9%; Score 102; DB 18; Length 466;
Best Local Similarity 23.3%; Pred. No. 0.11;
Matches 45; Conservative 35; Mismatches 75; Indels 38; Gaps 8;

QY 24 DRTLKAG-SDADGSDSDSLPHFRGKDSWIFYARREAAAILGLD---QKISHLTDD 78
Db 168 DRLVHPGWDQFQIDSITAPLPNARAKRDAMVDENEAQVLDVPTADRDLDATVAPE 227
QY 79 ELDALEDDVOKARAVRGLVENEED--SDSQSFPPVRKYMSNGT-----IDNFV-- 126
Db 228 EIEMEEDDISIAETERGVLDDHHYFSDSDSLPARKPLPKGTSTQSAWFIIDVSDS 287
QY 127 -----LDTAITMPTSRDGDGDLALFFGDKK-----SKQDASNVDEVELRQQ 169
Db 288 GSDTEEEBQDAMAM-DTAGNPEDG--VPPDRQDAMTEAGPSEYPOSEMFLDPSPE 343
QY 170 AQMEELAKTAKOE 182
Db 344 AQOLEYRASRRK 356

RESULT 6
US-09-733-089-17684
; Sequence 17684, Application US/09733089
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE OF INVENTION: Transcription In Plants
; FILE REFERENCE: 38-21(15300)D
; CURRENT APPLICATION NUMBER: US/09733,089
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 17684
; LENGTH: 845
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-733-089-17684

Query Match 10.6%; Score 99; DB 21; Length 845;
Best Local Similarity 23.7%; Pred. No. 0.54;
Matches 46; Conservative 32; Mismatches 66; Indels 50; Gaps 8;

QY 21 GRHDRTLKAGSDAGDSDSLPHFRGKDSWIFYARREAAAILGLQKISHLTDDDL 80
Db 519 GEIDHSTQK---EARMDSDCRRKHESGHDSS---SRVQOSSILSRNRLCHOLLEQC 572
QY 81 DAL-----FDDVQKARAV-----RRCGLVEDNEDSDQ 107
Db 573 DDLKYGSSNDFKAIKMKRLFEISVQLQEVPIQLPYASPLKSETNRLVQDGRNSCR 632
QY 108 SSFPV---RDKYMSNGTIDNFSLDTAITMPTSRDGDGDL-FFGDKK----SKDASNV 160
Db 633 NIDLOSDNDEDTFANVDNIGANTVTVLV-----DSDGDGDSVASFVDEKSSKSNANYI 688
QY 161 DVEELRQQAQME 174

```

```

Query Match      10.6%; Score 99; DB 21; Length 845;
Best Local Similarity 23.7%; Pred. No. 0.54;
Matches 46; Conservative 32; Mismatches 66; Indels 50; Caps 8;

QY   21 GRHDTLSKAGSDGDSRSDSPLPFRHGKDSWFWARREASAILGLDQIKSHLTDEL 80
     | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    519 GEIDHTQX---EARDMSDCVRRKHESGDHS--SRVQSQSILSRNRILCHOLLEQC 572
                                     ||| : ||| : ||| : ||| : ||| :
QY   81 DAL-----FDDVQKARAV-----RRGLVEDNESDSQ 107
     | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db   573 DDLKYGSSTNDFKAISMKRLEFISIVQLQEVPILPYASP LSKSSETNLRVQDGRNSSCR 632
                                     ||| : ||| : ||| : ||| : ||| :
QY  108 SSPPV---RDXYMNGTINDFSLDTAITMCTPRSDDDGDAL-FFGDKK---SKQDASNW 160
     | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   633 NIDLOSDNDIEDTFANVDNIGANTVVLV----DSDDGDSVASVFDEKSSDSKQNIYI 688
                                     ||| : ||| : ||| : ||| : ||| :
QY  161 DVEELRQQQAQME 174
```



```

; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 17565
; LENGTH: 775
; TYPE: prt
; ORGANISM: Oryza sativa
US-09-733-089-17565

```

[illegible]

```

RESULT 14
US-09-733-089-17566
; Sequence 17566, Application US/09733089
; GENERAL INFORMATION:
; APPLICANT: Dofson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(15300)D
; CURRENT APPLICATION NUMBER: US/09/733.089
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 17566
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-733-089-17566

```

```

Qy  108  SSFV---RRKYSNGTIDNFSLDTATIMPGTSPRSSDDGDAL-FFGDKK---SKQDASNV 160
      :      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db   663  NIILDSNDNEDYTFANVDNIGANTTVLV-----DSDGDGVASVFDKSSDSKQANAYI 718
      :      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy   161  DVEELRQQAQME 174
      :      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db   719  EESVLPQHAQQQE 732
      :      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 15
US-60-197-873-25538
; Sequence 25538, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejain, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81 US1 PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 25538
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 14
; OTHER INFORMATION: Xaa = Pro,Arg
US-60-197-873-25538

```

Search completed: April 25, 2001, 10:16:30
Job time: 331 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:32 ; Search time 7.94 Seconds
(without alignments)
72.702 Million cell updates/sec

Title: US-09-2335-416-1_COPY_602_784

Perfect score: 938
Sequence: 1 QEQSLRHVSQSLGSPAP.....ELRQQQAQMEALKTAKQEF 183

Scoring Table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 17899 seqs, 3154390 residues

Total number of hits satisfying chosen parameters: 17899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.5	8.2	1503	5	US-09-649-996-14
2	73.5	7.8	759	5	US-09-792-024-76
3	73	7.8	393	6	US-60-248-505-1134
4	72.5	7.7	425	6	US-60-248-505-737
5	69	7.4	1758	5	US-09-739-449-11627
6	68.5	7.3	694	5	US-09-739-449-9792
7	68	7.2	492	5	US-09-739-449-10209
8	68	7.2	693	5	US-09-336-910A-3
9	68	7.2	714	5	US-09-792-024-117
10	67	7.1	1042	5	US-09-792-024-106
11	66.5	7.1	529	5	US-09-739-449-8477
12	66.5	7.1	606	5	US-09-739-449-11845
13	66	7.0	684	5	US-09-739-449-12024
14	66	7.0	812	5	US-09-739-449-9492
15	65.5	7.0	727	5	US-09-822-246-2
16	65.5	7.0	727	5	US-09-822-246-4
17	65	6.9	422	5	US-09-781-417-85
18	64.5	6.9	395	5	US-09-739-449-11881
19	64.5	6.9	648	5	US-09-813-408-26
20	64.5	6.9	695	5	US-09-806-194-10
21	64.5	6.9	695	5	US-09-806-194-12
22	64.5	6.9	695	5	US-09-806-194-14
23	64.5	6.9	697	5	US-09-806-194-16
24	64.5	6.9	697	5	US-09-806-194-18
25	64.5	6.9	697	5	US-09-806-194-20
26	64.5	6.9	976	5	US-09-628-359-18
27	64	6.8	281	5	US-09-739-449-10656

28	64	6.8	319	5	US-09-739-449-12619	Sequence 12619, A
29	64	6.8	321	5	US-09-739-449-12453	Sequence 12453, A
30	64	6.8	435	5	US-09-739-449-8567	Sequence 8567, Ap
31	64	6.8	778	5	US-09-739-449-12243	Sequence 12243, A
32	63.5	6.8	220	1	PCT-US01-10484-78	Sequence 78, Appl
33	63.5	6.8	406	5	US-09-813-329-4	Sequence 4, Appl1
34	63.5	6.8	409	5	US-09-813-329-2	Sequence 2, Appl1
35	63.5	6.8	409	5	US-09-813-329-6	Sequence 6, Appl1
36	63.5	6.8	992	6	US-60-248-505-1112	Sequence 1112, Ap
37	63.5	6.8	1384	5	US-09-649-996-11	Sequence 11, Appl
38	63.5	6.8	3170	6	US-60-248-505-909	Sequence 909, App
39	62.5	6.7	110	1	PCT-US00-30036A-162	Sequence 162, App
40	62.5	6.7	110	1	PCT-US00-30036A-165	Sequence 165, App
41	62.5	6.7	112	1	PCT-US00-30036A-87	Sequence 87, Appl
42	62.5	6.7	112	1	PCT-US00-30036A-114	Sequence 114, App
43	62.5	6.7	331	5	US-09-739-449-9137	Sequence 9137, Ap
44	62.5	6.7	528	5	US-09-792-024-77	Sequence 77, Appl
45	62.5	6.7	587	5	US-09-781-417-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-649-996-14
; Sequence 14, Application US/09649996
; GENERAL INFORMATION:
; APPLICANT: Jono, Keith E.
; TITLE OF INVENTION: KINASE GENES AND USES
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/649,996
; FILING DATE: 29-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/976,255
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 229/182
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1503 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-649-996-14

Query Match 8.2%; Score 76.5; DB 5; Length 1503;
Best Local Similarity 24.7%; Pred. No. 7.3;


```
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11627
; LENGTH: 1758
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11627

Query Match          7.4%; Score 69; DB 5; Length 1758;
Best Local Similarity 21.2%; Pred. No. 44;
Matches 49; Conservative 32; Mismatches 76; Indels 74; Gaps 11;

QY 16 GSPAPGRHD-----RTLSKAGS-----DADGDSRSDSPLPHF-----RGKSDMFY 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 525 GRAAPGADIVFSWTERGIYRAGETVHAALARDVDGKAIEDLPLTFIFSRPDGVEDRREV 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 ARREA-----ASAILG-----LDQXI-----SHLTDDEL--ALFDDVOKA 90
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 585 SDGKALGGHADVPLQANAMRTWLTIRHTDPKTAATSEKSLVDVFPDRTFEDLSKA 644
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 91 RAVRRGLVEDNEDSDSQSPFVRDKYMSNGTIDNFSLDTAITMPTGTPRSDDDGDALFFG- 149
   : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 645 KOI-----DPGAETALDVGRLYGAPAGLTLGEITAIRPT--RTTDFEGYFGL 694
   : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 150 -DKSKQD-----ASNVDVEELRQOQAHEEALKTAQOE 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 ADEESEENRTTLADLPVLDEEGKASFNVOLDLPSTTQLLSANITVRMOE 745

RESULT 6
US-09-739-449-9792
; Sequence 9792, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9792
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9792

Query Match          7.3%; Score 68.5; DB 5; Length 694;
Best Local Similarity 25.9%; Pred. No. 15;
Matches 49; Conservative 25; Mismatches 74; Indels 41; Gaps 12;

QY 7 RHSVTNSQLGS---PAPGRHRTLSKAGSDAD---GDSRSDSPLPHFRGK-----50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 427 KHPVTRGLESGNEPPQWGRWERTISVDPPQGTVMGLGDQDPLLLNLRQEGRVAMLLS 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 51 DSDWFYARR--EASATILGLDOKISH--LTDDDELDAFDVQKARAVRGLVEDNE---DS 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 487 DQGLWARGFEGGPHVALYRIAHWLKKEPELE---EALTAIRANGTQITRQIGDD 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 105 DSQSS--FPVRDKYMSNGTIDNFSLDTAITMPTGTPRSD--DGDALF---FGDKSKQDA 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 544 PQQASVRFP-----SGTENAFPT--TEPGLYRIERRMDETGLFEIKNGDFTTLVHV 594
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 158 SNVDVEELR 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 595 GAVDAPEFK 603

RESULT 7
US-09-739-449-10209
; Sequence 10209, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 10209
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10209

Query Match          7.2%; Score 68; DB 5; Length 492;
Best Local Similarity 19.7%; Pred. No. 11;
Matches 35; Conservative 34; Mismatches 59; Indels 50; Gaps 9;

QY 20 PGRHRTLS-KAGSDADGDSR-----SDSPLPHFRCKSDMFYARREASAI-----65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 PGEEDQLVGKA-----GDEKTTVTTPADYPAANLAGKATFDVTVKVAAAANAVEIND 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 66 -----LGLD--QKISHLTDDDELDAFDVQKARAVRGLVEDNEDSDSQSPFVRDKYMS 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 ELAEKLGLESAEKLEIVKGQIESQFQGNVTRQVKRQ--ILDOLDENYKFDTPAG---LV 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 119 NGTIDNF--SLDTAITMPTGTPRSDDDGDALFFGDKKSKQDASNVDEELRQOQAHEE 174
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 DAEDFNWRQINTDLAQSGKTFAD-----TTEEAEREYRKLAEE 354

RESULT 8
US-09-336-910A-3
; Sequence 3, Application US/09336910A
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 3
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Marburg Virus
US-09-336-910A-3

Query Match          7.2%; Score 68; DB 5; Length 693;
Best Local Similarity 23.3%; Pred. No. 17;
Matches 27; Conservative 14; Mismatches 41; Indels 34; Gaps 3;

QY 68 LDKQISHLTDDDELDAFDVQKARAVRGLVEDNEDSDSQSPFVRDKYMSNGTIDNFSL 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 439 VDDKIEH-----ESTEDSSSSSFVDLNDPFFALLNEDEDTL 474
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:07:04 ; Search time 56.07 seconds
(without alignments)
799.286 Million cell updates/sec

Title: US-09-235-416-1
Perfect score: 4030
Sequence: 1 MSGGGTKVVVRPFNARE.....ELRQQQAQMEALKARQEF 784

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4030	100.0	784	20 Y06618	Thermomyces lanugi
2	1668.5	41.4	1816	21 B36227	Human kinesin-like
3	1658	41.1	1103	21 Y51328	Human KLIMP protei
4	1278.5	31.7	504	21 B63189	Gene 5 human secre
5	1276	31.7	503	21 B63190	Human secreted pro
6	814	20.2	955	15 R57365	K39 polypeptide of
7	814	20.2	955	17 W03691	Leishmania chagasi
8	662.5	16.4	2954	20 Y01632	Amino acid sequenc
9	649.5	16.1	975	19 W72745	Drosophila kinesin
10	625	15.5	411	19 W72745	Drosophila kinesin
11	625	15.5	441	19 W72744	Drosophila kinesin

12	584	14.5	1518	21 G40075	Arabidopsis thalia
13	584	14.5	1662	21 G31282	Arabidopsis thalia
14	580.5	14.4	1460	21 G40077	Arabidopsis thalia
15	580.5	14.4	1462	21 G40076	Arabidopsis thalia
16	580.5	14.4	1604	21 G31284	Arabidopsis thalia
17	580.5	14.4	1606	21 G31283	Arabidopsis thalia
18	554	13.7	1269	21 Y77955	A. thaliana enviro
19	542	13.4	469	21 G21665	Arabidopsis thalia
20	527.5	13.1	452	21 G21666	Arabidopsis thalia
21	504	12.5	398	21 G21667	Arabidopsis thalia
22	489	12.1	829	21 G31117	Arabidopsis thalia
23	489	12.1	834	21 G31116	Arabidopsis thalia
24	488.5	12.1	460	21 B56650	Human prostate can
25	471.5	11.7	1034	21 G31112	Arabidopsis thalia
26	471.5	11.7	1069	21 G31111	Arabidopsis thalia
27	471.5	11.7	1121	21 G31110	Arabidopsis thalia
28	464.5	11.5	324	19 W70235	Leishmania antigen
29	459.5	11.4	730	21 Y49949	Xenopus laevis kin
30	444	11.0	679	20 W88456	Human kinesin-rela
31	433.5	10.8	790	21 G41923	Arabidopsis thalia
32	433.5	10.8	794	21 G41922	Arabidopsis thalia
33	433.5	10.8	814	21 G41921	Arabidopsis thalia
34	403	10.0	147	21 G02949	Human secreted pro
35	381.5	9.5	726	21 G31118	Arabidopsis thalia
36	379.5	9.4	410	21 B56496	Human prostate can
37	366	9.1	959	21 Y77944	A. thaliana enviro
38	304.5	7.6	1201	20 W90345	Drosophila sp. Cos
39	302	7.5	154	21 B40661	Human cancer assoc
40	211.5	5.2	243	21 Y92345	Human ORFX ORF425
41	202	5.0	92	21 B41721	Human ORFX ORF1485
42	195	4.8	1612	16 R66457	Chimeric ALL-1/AF-
43	195	4.8	1612	18 W24094	Ras-binding protei
44	185.5	4.6	1829	20 Y07242	Actin-filament bin
45	155	3.8	752	18 W34178	Human transcriptio

ALIGNMENTS

RESULT 1
Y06618
ID Y06618 standard; Protein; 784 AA.
XX
XX Y06618;
AC Y06618;
XX
DT 26-OCT-1999 (first entry)
XX
XX Thermomyces lanuginosus kinesin motor protein TL-gamma.
DE
DE TL-gamma; kinesin; motor protein; microtubule; unc-104; infection;
KW neurodegenerative disease; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis.
XX
XX Thermomyces lanuginosus.
OS
OS
PN W09937659-A1.
XX
XX
XX 29-JUL-1999.
XX
XX 22-JAN-1999; 99WO-US01355.
XX
XX 23-JAN-1998; 98US-0072361.
PA (REGC) UNIV CALIFORNIA.
XX
XX Goldstein LSB, Sakowicz R;
XX
XX WPI: 1999-493950/41.
XX
XX N-PSDB; X87656.
XX
XX New nucleic acid encoding microtubule motor protein, used for
PT diagnosis of fungal infection and neurodegenerative disease

XX Claim 5; Page 70-71; 75pp; English.

XX This sequence represents Thermomyces lanuginosus TL-gamma, a novel

CC ATP-dependent, plus end-directed microtubule motor protein that is

CC a member of the unc-104 family and kinesin superfamily. The

CC invention provides TL-gamma nucleic acids (see X87656), proteins

CC and antibodies, and methods of screening for TL-gamma modulators

CC potentially useful for treating hyphal fungal infections and

CC diseases caused by mutated TL-gamma, e.g. neurodegeneration

CC involving anterograde axonal transport, such as Alzheimer's,

CC Parkinson's or Huntington's diseases or amyotrophic lateral

CC sclerosis. Detection of TL-gamma allows differentiation between

CC hyphal and non-hyphal fungal infections.

XX SQ Sequence 784 AA;

Query Match 100.0%; Score 4030; DB 20; Length 784;

Best Local Similarity 100.0%; Pred. No. 4e-300;

Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGGNIKVVVRFPFNAREIDRGAKCIVRMEGNOTILTPPPGAEKARKSGKTIMDGP 60

Db 1 msggnikvvvrpfnareidrgakcivrmegnotiltpppgaeekarksgktimdgpk 60

Qy 61 AFADRSYWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQGTGSGKSYMMG 120

Db 61 afadrsywsfdknapyaredlfqdlgvplldnafgynncifayggtgsgksysmmg 120

Qy 121 YGKEGVIPRICQDMFRINELQDKNLTCTVEYSLEYIYNERVRLNPNSTKGNLKVRE 180

Db 121 ygkegvipricqdmfrinelqdknltctveysleyiynervrldlnpstkgnlkvre 180

Qy 181 HPSGTPVEDLAKLVRSFQFIENLMDEGNKARTVAATNMNETSSRSASHVFTLTQKWH 240

Db 181 hpsgtpyvedlavlrvsfqfielnmdegnkartvaatnmnetsrshavftltqkwh 240

Qy 241 DEETKMDTEKVAKISLVDLAGSRATSTGATGARLKGEAENRSLSTIGRVIAALADMS 300

Db 241 deetkmdtekvakislvdlagseratstgatgarlkgeaenrslstlgrviaaaladms 300

Qy 301 GKQKNQILVPRDVLWLLKDSLGSSMTAMIAAISPADINFEETSLTRYADSAKR 360

Db 301 gkqknqilvprdvllwllkdslgssmtamiaaispadinfeetstltryadsakr 360

Qy 361 NHAVVNEDPNARMIREKEELAEALRSKIQSSGGGGAGGGPVESYPDTPLEKQIV 420

Db 361 nhavvnedpnarmirekeelaqlrsklqssggggagggpvessypdtplekqiv 420

Qy 421 SIQPDATVKKMSKAEIVQELNQSEKLYRDLNQWEEKLAKTERIHKEREAALEELGISI 480

Db 421 siqpdatvkkmskaeiveqlnqseklrldlnqwweeklakterihkereaaaleelgisi 480

Qy 481 EKGFGVGHKEMPHNLNLSDDPILAECLVYNIRPGQTRGVNVDQTAERLNGSKILK 540

Db 481 ekgfvgvghskemphnlslsddpilaeclyvynirpgqtrgvnvdqtaerlngskilk 540

Qy 541 EHCTFENVNDVVTVPNEKAAMVNGVRIDKPTRLSRYRIILGDFHFRNHPPEARAE 600

Db 541 ehctfenvndvvtvpnekaamvngvrldkptrlsryriilgdfhfrnhppeearae 600

Qy 601 RQEOSLLRHVTSNQLSPAGFRHRTLSKAGSDAGDSRSDSLPHFRGKDSDFYARR 660

Db 601 rqeosllrhvtsnqlspagfrhrtlskagsdagdsrdsdslphfrgkdsdfyarr 660

Qy 661 EASAILGLDQKISHLTDELDALFDVQKARAVRGLIVENEDSDSSFPVRDKYMSN 720

Db 661 easailglqdkishltdeldalfdvqkaravrrglivednedssdsgsfvprdkymn 720

Qy 721 GTIDNFSLDTAITMPGTPRSDDDDGALFFGDKSKQDASNVVDVEELRQOQAOEALKTA 780

Db 721 gtidnfsldtaitmpgtprsdgddgalfgdkskqdasnvvdveelrqoqaoealkta 780

Qy 781 KQEF 784

Db 781 kqef 784

RESULT 2

B36227 standard; Protein; 1816 AA.

XX B36227;

XX AC B36227;

XX AC B36227;

DT 19-FEB-2001 (first entry)

XX Human kinesin-like protein HKLP SEQ ID NO: 4.

XX Human; kinesin-like protein; HKLP; KIF1; cell division; cancer;

KW intracellular transport; neurological disorder; infertility;

KW biallelic marker; spontaneous abortion; neonatal chromosome disorder;

KW aneuploidy.

OS Homo sapiens.

PN W0200063375-A1.

XX 26-OCT-2000.

XX 20-APR-2000; 2000WO-IB00562.

XX 20-APR-1999; 99US-0130217.

XX (GEST) GENSET.

PA Bougueleret L, Dufaure-Gare I, Grel P;

PI WPI: 2000-665242/64.

DR N-PSDB; C66550.

XX An isolated or purified human kinesin-like protein (HKLP) encoding

PT polynucleotide used to detect HKLP polynucleotides in a sample

PT comprises a contiguous span of at least 12 nucleotides -

XX Claim 46; Page 189-192; 199pp; English.

XX The present invention describes the coding and protein sequences of the

CC human kinesin-like protein HKLP. It is thought that the protein could be

CC involved in neurological disorders, infertility, spontaneous abortion,

CC neonatal chromosome disorders, aneuploidy and cancers. This is due to its

CC function in the movement of microtubules. The protein shows homology to

CC the murine KIF1A and KIF1B proteins. The sequences disclosed in the

CC invention can be used in the isolation of similar human proteins and in

CC vector production. In addition, the biallelic markers shown can be used

CC in disease diagnosis and population studies.

XX SQ Sequence 1816 AA;

Query Match 41.4%; Score 1668.5; DB 21; Length 1816;

Best Local Similarity 45.0%; Pred. No. 1.2e-118;

Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;

Qy 4 GGNTKVVVVRFPFNAREIDRGAKCIVRMEGNOTILTPPPGAEKARKSGKTIMDGP 63

Db 3 gasvkvavrvfnrfsrskscskcliqnqgnstslinpkpkpke-----apksfs 51

Qy 64 FDRSYWSF-DKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGQGTGSGKSYMMG 122

Db 52 fdyswshstspedpcfasqnrnyndigkemllhafegynvcifayggtgagksytmmg 111

Qy 123 KEH--GVIPRICQDMFRINELQDKNLTCTVEYSLEYIYNERVRLNPNSTKGNLKVRE 180

Db 112 eesgagtiplqiceelfekind-nceemsysvevsymeicyervrldlnpknknlrvre 170

QY 181 HPSTGPVEDLAKLVRSFOEINLMDGNKARTVAATNMNETSSRSRAVFTLTQKWH 240
Db 171 hpllgpyvedlsklavsydydlmdagnkrtvaatnmnetsrshavftltvfkqh 230
QY 241 DEETKMDTEKAVAKISLVLDLAGSERATSTGATGARKKEAGAEINRSLSTGLRVAALADN-- 298
Db 231 dnetnlstekvskislvdlagseradstgatkgrikeganinksltlgkvisalaevdn 290
QY 299 ---SSGOKKQOLVPYRDSVLTWLLKDSLGNSMTAMTAATSPADINFEETLSTLRVADS 355
Db 291 ctskkskkkttdfipyrgsvlclwlrnlggnsrmtamvaalspadinydetlstrfydr 350
QY 356 AKRIKHAHVNEPDNARMIRELKEELAQLRSKLQSSGGG-----GGAG----- 399
Db 351 akqlknavinedpnaklvrelkeevtrikdliragglgdldldplldysgsgskylk 410
QY 400 -----GSGGPVEESYPDPDTPLEKQ-----IVSIQPDATVKK 431
Db 411 dfqnkhryllasengrphfstasmglss--pascslssqvgltsvtsig--erlmst 467
QY 432 MSKAEIVELNQSEKLYRDLNQTWEELAKTEIEHKEREAALEELGISIEK--GFVGPYH 489
Db 468 pgeeealerikeekliaelnetweeklrkteairmereallaemgvalredggtlgvfs 527
QY 490 SKEMPHLVNLSDPLLAELGVYNIKPGQTRVGNVNODTQAEIRLNGSKILKEHCTFENV- 548
Db 528 pkktphlvinedplmsecllyikdgttrvgadaerrqdivlsgahkeehcifrser 587
QY 549 ----DNVVTVPNEKAAMVNGVRIDKPTRLRSYRIILGDFHIFRPNHPEARAEQ 604
Db 588 snsgveivtlpepcerasyngkrsvqvlrgsnrlimgkhvfrfnhpeqararek- 646
QY 605 SLLRHSVNSQLSGPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREAS 664
Db 647 -----tpsaetpsepdvdfagrelle 668
QY 665 AILGLDQK-----TSHLTDDDELALDQ-----VQKARAVRGLVED 701
Db 669 k-qgidmkqemkrlqemellykkeeeadllleqgridyeklqalqkvetrslaet 727
QY 702 NEDSDQSSEFP 712
Db 728 teeeeeeerv 738
RESULT 3
Y51328
AC 51328
AC 51328
DT 17-APR-2000 (first entry)
XX Human KLIMP protein.
DE KLIMP: kinesin-like motor protein; cytotostatic; anticonvulsant; human;
KW anti-Alzheimer; anti-Parkinsonian; antidiabetic; anti-ulcerative; cancer;
KW immunomodulatory; anti-inflammatory; anti-AIDS; antirheumatic; treatment;
KW antiarthritic; diagnosis; neurological disorder; vesicular transport.
XX
OS Homo sapiens.
XX
XX US6013454-A.
XX
XX 11-JAN-2000.
XX 28-SEP-1998; 98US-0162373.
XX
XX 28-SEP-1998; 98US-0162373.
XX (INCY-) INCYTE PHARM INC.
XX
XX Tang YT, Corley NC, Patterson C, Guegler KJ;

XX WPI; 2000-126064/11.
DR N-PSDB; Z4744.
XX
PT Nucleic acid sequences encoding a human kinesin-like motor protein
PT (KLIMP) useful for the treatment of diseases associated with
PT inappropriate KLIMP expression such as cancers, neurological disorders
PT and disorders of vesicular transport -
XX Claim 1; Fig 1A-J; 38pp; English.
XX
CC This invention describes a novel human kinesin-like motor protein
CC (KLIMP) (I) which has cytostatic, anticonvulsant, anti-Alzheimer's,
CC anti-Parkinsonian, antidiabetic, anti-ulcerative, immunomodulatory,
CC anti-inflammatory, anti-AIDS, antirheumatic and antiarthritic activity.
CC (I) and the protein it encodes may be used in the prevention, treatment
CC and diagnosis of diseases associated with inappropriate KLIMP expression
CC such as cancers, neurological disorders and disorders of vesicular
CC transport. For example, (I) (and vectors containing (I) (iv)) and the
CC KLIMP polypeptide may be used to treat disorders associated with
CC decreased KLIMP expression such as cancers (e.g. lymphoma, melanoma and
CC cancers of the breast lung and prostate), neurological disorders (e.g.
CC epilepsy, Alzheimer's disease and Parkinson's disease), disorders of
CC vesicular transport (e.g. diabetes mellitus/insipidus, Grave's disease
CC and gastric/duodenal ulcers), and some immune/inflammatory diseases (e.g.
CC acquired immune deficiency syndrome AIDS), rheumatoid arthritis and toxic
CC shock syndrome). This sequence represents the human KLIMP protein
CC described in the method of the invention.
XX Sequence 1103 AA;
Query Match 41.1%; Score 1658; DB 21; Length 1103;
Best Local Similarity 47.3%; Pred. No. 3.7e-118;
Matches 333; Conservative 120; Mismatches 159; Indels 114; Gaps 16;
QY 4 GGNIKVVRVRFNAREIDRGAKICIVRMENQOTILTPPGAEEKAKSGKTMIDGPKAFA 63
Db 3 gasvkvavrvrfnareidsqdkcvvmgmgnttslnp-----kqskdapsft 51
QY 64 FDRSYWSFDKNA-PNYARQEDLFODLGVPLLDNAFKGYNNCIFAYGOTGSGKSYSMGYG 122
Db 52 fdysywshtstedpqfasqgyrldigeemllhafegynvcifaygqtagksytmgrq 111
QY 123 K--EHGVIPRQDMFRINELKOKNLTCTVVSYLEIYNERVDRLLNPSTKGNLKVR 180
Db 112 epqgggvlpqclcdlfrvsenq-saqlsyvsyvsymeicycervrdllnpksrgslrvre 170
QY 181 HPSTGPVEDLAKLVRSFOEINLMDGNKARTVAATNMNETSSRSRAVFTLTQKWH 240
Db 171 hpllgpyvqdlslavtsyadiadlmdcgnkartvaatnmnetsrshavftltvfkqh 230
QY 241 DEETKMDTEKAVAKISLVLDLAGSERATSTGATGARKKEAGAEINRSLSTGLRVAALADMS 300
Db 231 dqltidsekvsislvdlagseradsgarmglkeganinksltlgkvisaladms 290
QY 301 GKOKKQOLVPYRDSVLTWLLKDSLGNSMTAMIAAISPADINFEETLSTLRVADSARKIK 360
Db 291 -kkrksdfipyrdsvltelkenlggnsrmtamiaalspadinyeetlstrlyadrktqir 349
QY 361 NHAVVNEPDNARMIRELKEELAOLRSKLQSSGGGGAGG----- 400
Db 350 cnaignedpnarlirelqeavarlirelmaqglasaleglkteegsvrgalpasvappa 409
QY 401 -----SGGPVEESYPDPDTPLEKQIVSIQPDATVKKMSKAETVQLNOSKLYRDL 451
Db 410 pvaspsptthgelepfspst--esql-----gpeeamerlqetekiael 454
QY 452 NQTWBESKLAKTEIEHKEREAALEELGISIEK--GFVGPYHSEMPHVLNLSDDPLAECL 509
Db 455 netweeklrkteairmereallaemgavredggtvgvfpakktphlvlnedplmsecl 514
QY 510 VYNIKPGQTRVGNVNODTQAEIRLNGSKILKEHCTFENV-----DNVVTVPNEKAAMV 564

Db 515 lyhikdgtrvvgvdm-----ikltgqfrehclfrsibpdpdevvvtclepcegaetyv 570
 QY 565 NGVRIDPTRLRSYRIILGDFHIFRHNHDEEAREEQSLLRHSVTNSQLGSPAPGRH 624
 Db 571 ngkivteplvksnrlvmgknhvfrfnhpeqarler-----gypvp--- 614
 QY 625 DRTLKAGSDAGDSRSLPHPRGKDSOWFYARREASAILGDKIHLTDDELAL 684
 Db 615 -----ppppsepdwnfaqkelleg-qgidikle-mekriqdl 650
 QY 685 FDDVQKARVRRGLVEDNE---SDS 707
 Db 651 engyrkekeeadlllegqriyadsds 676

RESULT 4
 B63189
 ID B63189 standard; Protein: 504 AA.
 AC B63189;
 DT 26-MAR-2001 (first entry)
 XX Gene 5 human secreted protein homologous amino acid sequence #115.
 XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; neoprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;
 KW food additive; preservative.
 XX Drosophila melanogaster.
 XX W0200061629-A1.
 XX 19-OCT-2000.
 XX 06-APR-2000; 2000MO-US09071.
 XX 09-APR-1999; 99US-0128694.
 PR 20-JAN-2000; 2000US-0176931.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX Ruben SM, Komatsoulis G;
 PI WPI; 2000-647420/62.
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 used in preventing, treating or ameliorating a medical condition -
 PS Disclosure; Page 487-488; 533pp; English.

XX F22373 to F22421 encode the human secreted proteins given in B63134 to
 CC B63182. B63183 to B63231 represent more human secreted proteins and
 CC polypeptides homologous to them. Human secreted proteins have activities
 CC based on the tissues and cells the genes are expressed in. Examples of
 CC activities include: immunosuppressive; antiarthritic; antirheumatic;
 CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 CC neoprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also
 CC used in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders

CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. F22364 to F22372 and B63133
 CC represent sequences used in the exemplification of the present invention.
 XX

SQ Sequence 504 AA;

Query Match 31.7%; Score 1278.5; DB 21; Length 504;
 Best Local Similarity 50.9%; Pred. No. 1.3e-89;
 Matches 274; Conservative 77; Mismatches 144; Indels 43; Gaps 6;

QY 59 PKAFAPDRSYWSDKNAPVARQEDLFDQDLGVPLDFAFGYNNCINFCAYGOTSGKSYSM 118
 Db 8 pktfahdchfyslmpedenfasqetvfdcvgrgildnafgynacifaygqcgsgsksym 67
 QY 119 MGXGKEHGVIPRICQDMFRRLNELQKDKNLCTCTVEVSELYEINERVDLLNPS-TKGNLK 177
 Db 68 mgtqeskgilprlclqifsaian-kstpelmykvevsymelynekvhdllqpkngsklk 126
 QY 178 VREHPTGTPVEDLAKLVVRSFQIENLMDEGNKARVATNMNETSSRSHAVTILITQ 237
 Db 127 vrehnmgpyvvglsqilavtsyqgidlnlmetegnksrtvaatnmnaesrshavsvltq 186
 QY 238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLTLGRVIAALAD 297
 Db 187 iltldqatgvsgkevsmldvldagseravktgavgdrlkegsnknsltlglvsklad 246
 QY 298 MSSGKQKKN-QLVPYRDSVLTWLKDSLGNSMTAMIAAISPADINFEETLSTLYADSA 356
 Db 247 qsgngksgndkfpyrdsvltwlkdnlgnsrtvmvatispsadnyeetlstlryadra 306
 QY 357 KRIKNHAVNEDPNARMIRELKEALQIRSKLQSSGGGGGAGGGGGGPPVESSYPPDPLE 416
 Db 307 krivnhavvnedpnarilirelhevetrismkha-----tgspv----- 346
 QY 417 KQIVSIQOPDATVKKMSKAEIVOLNOSKLYRDLNLTWEEKLAKTEIHKEREAALEEL 476
 Db 347 -----gdvqdklaesenlmkqslvweeklvkterlqnergqalekm 388
 QY 477 GISIEKGFVGPYHSEKMPHLVNLSDPDLAECLVYNKPGQTRVGNVNDQTQABIRLNGS 536
 Db 389 gisvqas--gikveknkylvlnlnadpslnellylkdrtliggrtisgqpdqisgl 446
 QY 537 KILKEHCTFENVNDNVTVIPNEKAAMVNGVGRIDKPTLRSGYRIILGDFHIFRHNHP 594
 Db 447 giqpehcvittedsglymepvgqarcfvngsaavektpqlngdrilwgnhhffrvnsp 504

RESULT 5
 B63190
 ID B63190 standard; Protein: 503 AA.
 AC B63190;
 DT 26-MAR-2001 (first entry)
 XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:116.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; neoprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing; skin aging;

Qy	291	VTAALDMSSGKOKNQVLVPYROSVLTWLLKDSLGGNSMTAMIAALSPADINFEETLSTL	350
Db	244	visaladqsqag--snkfkpyrdsvlwtalkdsglsgnskamvatvспаадныdetlsl	302
Qy	351	RYADSAKRIRKNHAVNEDPNARMIRELKEELAQLRSKLQSGGGGGAGGGGGVPEESYP	410
Db	303	ryadrakrhvnhavnedpnarirdlireevkeqlrta-----	343
Qy	411	PDTPLEKQIVSYQQPDATVKKMSKAETVQLNQSEKLYRDLNWTPEKLAKTEIHWERE	470
Db	344	-----eamspkeldrleeseklqemvtweeklrkteelagrq	384
Qy	471	AALBELGISTEKGFCYHSEMPHVLNLSDDPLAELCLVLYNIKPGQTRVGNVQDTPQAE	530
Db	385	kqlseqlgislqas--gikvgdkcflvlnadpalnellyyike-htliglsansq---d	438
Qy	531	ITLNGSKILKECHCTFE-NVDNVVTIYPVNEKAAMVWNGVRDKPTRLRBSYRIILIGDHPHF	589
Db	439	iqicymglpheiichliditsegvmltpqkntftfngdssvsspiqlhngdrilwgnnhff	498
Qy	590	RFNHP	594
Db	499	rlo1p	503

DR	WPI; 1994-249402/30.
XX	N-PSDB; Q70152.
PT	
PT	Diagnosis of Leishmaniasis - by determining the presence of antibodies that bind to a K39 repeat unit antigen
XX	
PS	Disclosure; Page 12-15; 28pp; English.
XX	
CC	The K39 polypeptide comprises a number of repeated units (described in R57366). Detection of antibodies directed against this repeated unit in a patients sample is indicative of leishmaniasis. The antigenic repeat unit can itself be used as a vaccine to protect against infection by a leishmania parasite.
XX	
SQ	Sequence 955 AA;
Query Match	20.2%; Score 814; DB 15; Length 955;
Best Local Similarity	28.1%; Pred. No. 1.4e-53;
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 1	

```
QY 7 IKVVVRVPFNARE--IDRGAKCIVRMENOTILTPP-----PGAEEKARKSGKTIMD 57
DB 13 VKSVVRVPLNERENNAPGCTKTVAAKQAQAAVTVKVLGSSNNSGAAESNGTARRVAQD 72
QY 58 GKAFADRSYWSF---DKNAPNTARQEDLFDQLGVPLLDNAFKYNNCIFAYGOTSGK 114
DB 73 ---fqdhvfwsvetpdacgatpatqadvfrtgyplvqhafgdfnscifayqgtsgk 128
QY 115 SYSWMG-----YKKEHGVIPRICODMFRRLNELOKKNLTCTVEVSYLETYNVRDLL 168
DB 129 tytmgadvlsalsgegnvtpiclleifarkasvafdrskitfildslggnsktmiat 188
QY 169 NPSKGG-----NLKVRHPSTGPPVEDLAKLVRSFQEIENLMDGKNKARTVAATNM 220
DB 189 gkrkkgvkgggeeeyvdrhpsrgvlegqrlyevslddvrlieigngvrhtastkm 248
QY 221 NETSRSHAVFTLT---TQKWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLK 276
DB 249 ndrshrshaimllireertmttksgetirtagkssrmlvdlagservagsgvqqfk 308
QY 277 EGAETNRSLSLGRVIAALADMS--GKQKNQOLVPYRDSVLTLLKSLGNSMTAMTAA 335
DB 309 eathinlsittlgrvidladmatkagavsvafdrskitfildslggnsktmiat 368
QY 336 ISPADINFEETLSLRVADSARKIKNHAVNEDPNARIRELKEELAOGLSKLSSGGGG 395
DB 369 vspalnyeetlstrysardinvavqnedprarrireeqmedmrqam-----421
QY 396 GGAGSGGPVEE-----SYPPDTPLEK 417
DB 422 --agdpayvselkkkllallesaqkraadlqalererhngvqerllrateaekseles 479
QY 418 QIVSIQ-----PDATVKMKSAEIVEQL 441
DB 480 raalqleemtatrrqadkmqalnrlrkeegarkerellkemakdaalskvrirkdaeia 539
QY 442 NQSEKLYRDLNQTWEELAKTEIHKREAALEELGISIEKGFVGPYHSEKEMPHLVNLS 501
DB 540 serek-----estvaqlereqrerevaldal-----qthqrklqealesse 581
QY 502 -----DPLLAELVNIKPGQTRGVNQNQDAEIRLNGSKILKEHCTFENVDNVVTV 555
DB 582 rtaaerdqlq--ltelqsertqlsvtdrerltr-dlqiryegetelardvalca 638
QY 556 PNE-----KAAMVNGVRIDKPTBLRSYRIILGDFHFRFNHPEEAERQESLRLHS 610
DB 639 aqenearhyaavfhlqtlllelatewedalr-----eralaeadeaaaae-- 682
QY 611 VTNQLGSPAPGRHDTLSKAGSDADGDSRSDSPLPHFRGKDSDFVYARREAAAILGLD 670
DB 683 -----ldaaastsqn-----aresaceritsle 705
QY 671 QKISHLTDELDFDDVQKARAVRGLIVEDNEDSDSSPPVRDKYMSNCTIDNFSLD 730
DB 706 qqlre-seeraaelasqlaataakssaeqndrentlatleqires-----earaael 757
QY 731 AITWPGTPRSDDDGALFFGDKKSKQDASV-----DVEELRQQQAQMEELAKTAK 781
DB 758 asqlaataa-----kmsaeqndrentlatleqirdseeraaelasqllesttaak 807
```

RESULT 7

W03691 ID W03691 standard; Protein; 955 AA.

XX AC W03691;

XX AC W03691;

DT 09-MAR-1997 (first entry)

XX Leishmania chagasi K39 antigen.

XX Leishmania chagasi; acidic ribosomal antigen; LcP0;

XX Leishmania chagasi; K39 epitope; K39.

```
XX OS Leishmania chagasi.
XX PN W09633414-A2.
XX PD 24-OCT-1996.
XX PF 19-APR-1996; 96WO-US05472.
XX PR 21-APR-1995; 95US-0428414.
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG;
XX DR WPI; 1996-485884/48.
XX DR N-PSDB; T42166.
XX PT New Leishmania acidic ribosomal P-protein family poly:peptide - used
XX PT to develop prods. for diagnosis, detection and protection against
XX PT Leishmania infections
XX PS Disclosure; Page 36-43; 76pp; English.
XX CC Compounds including polypeptides that contain at least an epitope of
XX CC the L. chagasi acidic ribosomal antigen LcP0 are useful in a variety
XX CC of immunoassays for detecting Leishmania infection. Portions of
XX CC LcP0 (T42164) contg. at least the 17 C-terminal amino acids (T42165)
XX CC have been found to generate a signal in an ELISA that is equivalent
XX CC to that generated by the full length LcP0. A combination
XX CC polypeptide may also be used, comprising an LcP0 epitope along with
XX CC an epitope derived from the Leishmania K39 antigen (T42166), pref.
XX CC the K39 repeat unit antigen having the sequence given in W03690.
XX SQ Sequence 955 AA;
```

Query Match 20.2%; Score 814; DB 17; Length 955;

Best Local Similarity 28.1%; Pred. No. 1.4e-53;

Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

```
QY 7 IKVVVRVPFNARE--IDRGAKCIVRMENOTILTPP-----PGAEEKARKSGKTIMD 57
DB 13 VKSVVRVPLNERENNAPGCTKTVAAKQAQAAVTVKVLGSSNNSGAAESNGTARRVAQD 72
QY 58 GKAFADRSYWSF---DKNAPNTARQEDLFDQLGVPLLDNAFKYNNCIFAYGOTSGK 114
DB 73 ---fqdhvfwsvetpdacgatpatqadvfrtgyplvqhafgdfnscifayqgtsgk 128
QY 115 SYSWMG-----YKKEHGVIPRICODMFRRLNELOKKNLTCTVEVSYLETYNVRDLL 168
DB 129 tytmgadvlsalsgegnvtpiclleifarkasvafdrskitfildslggnsktmiat 188
QY 169 NPSKGG-----NLKVRHPSTGPPVEDLAKLVRSFQEIENLMDGKNKARTVAATNM 220
DB 189 gkrkkgvkgggeeeyvdrhpsrgvlegqrlyevslddvrlieigngvrhtastkm 248
QY 221 NETSRSHAVFTLT---TQKWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLK 276
DB 249 ndrshrshaimllireertmttksgetirtagkssrmlvdlagservagsgvqqfk 308
QY 277 EGAETNRSLSLGRVIAALADMS--GKQKNQOLVPYRDSVLTLLKSLGNSMTAMTAA 335
DB 309 eathinlsittlgrvidladmatkagavsvafdrskitfildslggnsktmiat 368
QY 336 ISPADINFEETLSLRVADSARKIKNHAVNEDPNARIRELKEELAOGLSKLSSGGGG 395
DB 369 vspalnyeetlstrysardinvavqnedprarrireeqmedmrqam-----421
QY 396 GGAGSGGPVEE-----SYPPDTPLEK 417
DB 422 --agdpayvselkkkllallesaqkraadlqalererhngvqerllrateaekseles 479
```

Qy	418	QTVSI00-----	-----PDTVKKMSKAIEVQOL	441
Db	480	raaalqemlatrrqdqmqlnlrlkeeqarkerellkemmaakdaelskvrrrkdaeta	539	
Qy	442	NOSCKLYRDUNOTWEKLAKTEIHXERAAALELGISIEKGFGVGYHSKEMPHLVNLSLSD	501	
Db	540	serekl-----estvaqlreqrearevaldal-----qthqrklqealesse	581	
Qy	502	-----DPLLAELCVNYNKPGOTRVGNQNODQAIRLNGSKTLKEHCTFENVNVVTIV	555	
Db	582	rtaaerdqllqq--itelqserqlsqvvtddrnltr-dlgrlqeygetelardvalca	638	
Qy	556	PNE-----KAANWVCVRIDKPTRLRKSQYRIILGDHFIFRNHPPEARAEQRQESLLRHIS	610	
Db	639	aqemearyhaavfhlgltllelalewedarl-----eralaeardeaaaae---	682	
Qy	611	VTSQLGSPAPGRDHRTLSKAGSDADGDSRSDSLPHFRGKDSDFWFAARAAASAILGLD	670	
Db	683	-----ldaaaastsgo-----aresacerltsle	705	
Qy	671	OKISHLTDDDELDAFDVQKARAVRRGLVEDNESDSOSSFPVPDYKYMSGTINDISLOT	730	
Db	706	qtlre-seeraaelasqlataakaasaacqdrntatleqlfres-----earaael	757	
Qy	731	AITWPGTPRSDDDGALFFGPKKSODASNv-----DVUELROQQOAMQEALKTKAT	781	
Db	758	asqlleaLaaz-----kmaaedcndrentatlleqldrseeraaelasqllesttaak	807	

RESULT	8
Y01632	
ID	Y01632 standard; Protein; 2954 AA.
XX	
AC	Y01632;
XX	
DT	22-JUN-1999 (first entry)
XX	
DE	XX
DE	Amino acid sequence of centromere-associated protein-E (CENP-E).

CENP-E; centromere-associated protein-E; ATPase activity;
plus end-directed microtubule motor activity; chromosome congression;
microtubule binding activity; chromosome movement; mitosis;
cell proliferation; tumor; metastasis; vascular malfunction;
inflammatory disease; immune disease; angiogenesis; hypertension;
restenosis; fungal infection; selective herbicide; fungicide;
resistance; plant growth regulator; activator; cancer cell marker.

of chromosomes during mitosis. Modulators of CENP-E can thus control cell proliferation. Agents that modulate CENP-E activity are lead therapeutic, biotechnological and diagnostic agents, e.g. for treatment of unwanted cell proliferation (typical of many examples are tumors and metastases; vascular malfunction; inflammatory and immune diseases; angiogenesis; hypertension; restenosis; and fungal infections), also as plant protection agents (selective herbicides, fungicides and insecticides) and plant growth regulators or activators for improving yields. CENP-E is also a diagnostic marker for dividing cells, including cancer cells.

Query Match 16.4%; Score 662.5; DB 20; Length 2954;
Best Local Similarity 26.8%; Pred. No. 3.1e-41;
Matches 228; Conservative 147; Mismatches 249; Indels 227; Gaps 29;

QY	1	MSGGNIKVVVRPFNAEIRGARCIIVRMEGNOTILTPPPGAEBEAKRSGKGTIMDGP	60
Db	1	msgdavagvavrplqre-----ggdqa1-----qwkagntisqvdgk	43
QY	61	AFADRSYFSFKNAPNARQEDLFODLGVPLLDNAPKGYNNCIFAYGOTGSGKSYSMG	120
Db	44	sfndrvfnshests-----g1ygeiavpiirsalqgngt1fayqtsagkctymmg	96
QY	121	YKGEHGVFIPRICDMFRINELQKKNLTCTVEVSYLEIYINERVDLL--NPSTKGNLKYR	179
Db	97	tpnslg1ipqalqevfk1qei---poref1lrvsyemelynetvkdllcdrrckpleir	153
QY	180	EHPSTGPGYVEDLAKIVRSPOETENLMDSGNKARTVAATNMNETSSRSHAVETLTTQK	239
Db	154	edforvnyadlteelvmvpehviqwkkeknrhyetkmndhsrshftlfrmvcsrd	213
QY	240	HDEBTKMD-----TERKAKISLVDLAGSRATSTGATGARKLKEGAEINSLSTLGRVIAAL	295
Db	214	rndptnscndgavmshlnvldagseaaqdaegvrlkegcnnlrsflilgqvkk1	273
QY	296	ADMSGKQKNOLVPRDSVLTWLLKXDSLGNSTMTAMAAISPADINFEETSLSTLAYADS	355
Db	274	sdgqag-----gfnyrdsktlr1qns1lgnaktviictip--vstde1st1lqfast	326
QY	356	AKR1KNHVVNE--DPNARMTRELKEELAQLRSKLQSGGGGGGGAGGSGGVPESVPPDT	413
Db	327	akhrntphnvne1ldeael1kryrke1l-----	354
QY	414	PLEKQIVSIQOPDAT--VKKSKAE1VEQUNOSEKLYLDQLNQTWEEKLAKTEIHKREAA	472
Db	355	d1kkq1en1lesstekaamak-----htql1aelkqlhkeredr	396
QY	473	LEE1LG1SI-----EKGFVGPYHSEKEMPHVLNLSDDPLLAELVYNIKGP	516
Db	397	1wh1n1lvassqesqdgqrkrkrvtwppq1qns1hasgvsdfm1sr1-----	448
QY	517	QTRGVNNDQTAER1LNGSKILKEHCT--FENVDNVVTVIPNKEAAVMVNGVRIDKPTRL	575
Db	449	pgnfsskakfsmfsspe1ddsvctefsd1lsmnds-----ng--1daewn1	496
QY	576	RSGYR1ILGDH1F1FNHPPEAEARQEOSLLRHSVNSQLGSPAGPHDRTLSKAGSDA	635
Db	497	as-----kvth-----reks1hqem1ofg1-----	518
QY	636	DGDSRSDSPLPHFRGKSDSMFYARREAA5AILGLDOK1SHLTDDELALFDVQOKAVR	695
Db	519	sdsvqfndsskenq1qy1pkd--sgdmaecrkasf--ekelt1lqq1lqskseek	569
QY	696	RLVEDNE---DSDSSGSPFPKDYKMSNGTIDNFS1D1AT1MPTCTPSSDDGDGAL---	747
Db	570	kelvqsfelkiealeeq1skvkanlwmtnsrsh-snaev-----q1dvekevvrkem	622
QY	748	FFGDKKSKQDASNDVDEE-----LROQQAAMEAL-----	777
Db	623	svldg---sagynasnd1qdsdvdk1lsshdce1hrkmlcek1vd1leef1en1lkkse	680

QY 778 ----KTAKQEF 784
Db 681 ndkqkseqdfl 691

RESULT 9

ID W72746 standard; Protein; 975 AA.

AC W72746;

DT 11-JAN-1999 (first entry)

DE Drosophila kinesin.

XX Drosophila; kinesin; separation; hybridisation; target site;
KW complex mixture; motor protein; actively transported; separated;
KW microtubule.

OS Drosophila sp.

XX Key Location/Qualifiers

FT Misc-difference 557

FT /note= "encoded by AAT"

XX US5830659-A.

XX 03-NOV-1998.

XX 13-SEP-1996; 96US-0713815.

XX 13-SEP-1996; 96US-0713815.

XX (UTAH) UNIV UTAH RES FOUND.

XX Stewart RJ;

XX WPI; 1998-609236/51.

DR N-PSDB; V671b2.

XX Separation of selected molecules, e.g. DNA, from complex mixtures -
PT uses specific apparatus to allow the selected molecule to bind to
PT motor proteins, and be actively transported and separated away along
PT microtubules

XX Disclosure; Column 17-24; 24pp; English.

XX A method has been developed of separating a selected molecule from a
CC mixture of molecules. The method comprises: (a) a separation device
CC comprising a loading reservoir and a receiving reservoir coupled by a
CC channel with microtubules immobilised on its surface and aligned parallel
CC to a longitudinal axis of the channel; (b) loading the loading reservoir
CC with an aqueous solution of the mixture of molecules; (c) adding a motor-
CC ligand composition and ATP to the solution, where the motor-ligand
CC comprises, (i) a processive motor capable of attaching to the immobilised
CC microtubules, and moving in the presence of ATP as source of chemical
CC energy, and (ii) a ligand coupled to the motor protein, where the ligand
CC is capable of binding the selected molecule, so that the ligand binds the
CC selected molecule and the motor protein attaches to the immobilised
CC microtubules and transports the bound selected molecules along the
CC receiving reservoir; and (d) removing the selected molecule from the
CC receiving reservoir. The method and the system are used for the
CC separation of specific molecules from complex mixtures. The molecule to
CC be separated (e.g. DNA) binds to the motor protein due to the presence of
CC the specific binding ligands. Activation of these enables them to travel
CC down a preformed channel in a specially made piece of apparatus. They can
CC then be removed easily without contaminants of other mixture particles.
CC The present sequence represents Drosophila kinesin from the present
CC invention.

XX Sequence 975 AA;

Query Match 16.1%; Score 649.5; DB 19; Length 975;
Best Local Similarity 27.4%; Pred. No. 5.6e-41;
Matches 232; Conservative 136; Mismatches 282; Indels 197; Gaps 32;

QY 6 NIKVVRVRFNAREIDRGAKCIVRMENGTILTPPPGAEEKARKSKOKTMDGPKAFAD 65
Db 12 slkvcrfrplndseekagskfvvkf-----pnnveenc-----islagkvylfd 56
QY 66 RSYWSFDKNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGOTGSGKSYSM---MGYG 122
Db 57 kvf-----kpn-asqekvyneaksivtdlagyngdtifayggtsgkhtmegvlgds 109
QY 123 KEHGVIPRIQODMRRINELQKOKNLCTCTVEVSLEYINERVDLLNPSTKGNLKVREHP 182
Db 110 vkqgiiprivndifnhiyam--evnlefihikvsvyeylmdkirdldvs-kvnlsvhekd 166
QY 183 STGPYEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSSHAVFTLTQKWHDE 242
Db 167 nrpyvkgaterfvsdpdvfeleegksnrhnavtmnnehsrshsvflinvkqenlen 226
QY 243 ETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTCGRVLAALADSSGK 302
Db 227 qkkl-----gklylvdlagsekvsktgaegtvideakninksalsalad---gn 279
QY 303 QKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAATSPADINFEETLSTLYADSAKR1KNH 362
Db 280 kth---lpyrdskltrilqeslgnarttivccpsafnesetckstldfgraktvknv 336
QY 363 AVNE-----DPNARM---IRELKEELAQLR-----ALEELGI 478
Db 337 cvneeltaeewkryekekknarklgkvekielelarwagetvkaeeqinmedlmea 396
QY 387 -----KLQSSGGGGGAG-----GSGGPVESYPPDTPLEKQIVSIQPDATVK 430
Db 397 stpnleveaaqtaaaalaqaqratalanmsasvavneqarlaterlyqqlddkdein 456
QY 431 KMSK--AEIIVEQLNQSEKLYRDLNQTWEKIAKTEETHKREA-----ALEELGI 478
Db 457 qdsyaeqikeeqvmeqeellianarreyetlqsemariqqenesakeevkvlqaleltv 516
QY 479 SI-----EKGFVGPYHSKEMPHLVNLS--DDPLLAECV----- 510
Db 517 nydksqeidnknkdidalneelqkqsvfnaastelqqlkdmsshqkkritemltnlr 576
QY 511 -----YNIRPGQ-----TRGVNVNQD--TQAEIRLNGSKILKEHCTFENVDNV 551
Db 577 dlgevqaiapgessidlkmalsagtdaskveedftma--rlfiskmkte-----akni 628
QY 552 VTIVPNEKAAVMVNGVRIDKPTRLSGVRIILGDFHFRFNHPEARAERQEQSLL----- 607
Db 629 agrcsnmnetqadnknkl-----seyekdlgeyrilishq--earmkslqesmraen 679
QY 608 RHSVNTNSQLSPAPGRIDRTLSKAGS-----DADGDSRSDSPLPFRGKDSDFYARREA 662
Db 680 kkrileeqidsl---reecaklkaahsvavnaeekgraeelrsmfidsqmd-----lrea 732
QY 663 ASAILGLDQKISHUTDD-----ELDALFDDVQKARAVRRLGLVEDNE---DSDSQSFPV 713
Db 733 ht-----rqvselrdeilaakqhemdemkdvhqklliahqgmtdayekvrqedakssel 786
QY 714 RDKYMSN 720
Db 787 qniilt 793

RESULT 10

W72745

ID W72745 standard; protein; 411 AA.

XX W72745;

XX 11-JAN-1999 (first entry)

Db 637 v--ddpehl 643

RESULT 13

G31282

ID G31282 standard; Protein; 1662 AA.

XX AC G31282;

XX DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 37541.

DE Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 06-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148585.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159684.	
PR	18-OCT-1999;	99US-0159588.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			14.5%; Score 584; DB 21; Length 1662;
Best Local Similarity			29.3%; Pred. No. 1,3e-35;
Matches 196; Conservative 102; Mismatches 195; Indels 176; Gaps 27;			
Qy	17 NAREIDRAKAC-----IVRMEGNQTLPPPGAEERKARKSGKTIIMDGKAFADRSYWSFD	72	
Db	41 nappidrtntstphrsmrmk-npprrpppsnplkrkisaetates-----gfsds-----	90	
Qy	73 KNAPYARQEDFDQDGLVPLLDNAFKYNNCIFAYGOTGCKSYSMWG-----YG	122	
Db	91 -----gvkeqmqqlvsgaplvenclsgfnssvraygqcgsgktyimmgpanglleehlcg	144	
Qy	123 KEHGVIPIRCODMFRINE-----LOKOKNLCTCTVEVSYLEIYNERNVRLN-	169	

Db	145 dqrgltpvrfarlikevmpgfvqfvvltlnksmllkn-sitnvghysrfsdwlhy	203	
Qy	170 -----PSTKGNL-----KVREHPSTGYPVEDLAKLVVRSFQOE	201	
Db	204 isiddlqrannrptgpkpdkghvhrfslmchqiredvkgvyvenlteeyvknld	263	
Qy	202 IENLMDEGNKARTVAATNMNETSSRSHAVFTLTITOKWHEETKMDTEKAKISLYDLAG	261	
Db	264 vsqlllkglnrrtgatvntesshsvfcvvesrcknvadgissfktsrlnlvdag	323	
Qy	262 SERATSTGATGARLKEGAENRSLSTPLGRVIAALADMS-SGOKKNQLVPYRDSVLTWLL	320	
Db	324 serqktgaagerlkeagnlnrsalsqnlilnlaeisqtgkprh---lpyrdarltfl	380	
Qy	321 KDSLGSNSMTAMIAISPADINFEETLSTLRVADSAKRIKNHAVVNE----DPN--ARMI	374	
Db	381 qeslgnaklanvcavspqrs--etfstlrfagrakalnkvavnevmqddvnlrgvl	438	
Qy	375 RELKEELAOLRSK-----LQSSGGGGGAGSGGPVEESYPP	411	
Db	439 hqlrdelqrmkndgnptnppvaystawnrrslnlrsfgl-----hpr	484	
Qy	412 DTPLKQ--IVSIOQPDATVKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEETHKER	469	
Db	485 slphedndgdiemeideavaerlcvgqlqslaseginhdmm-----rvkslhasd	536	
Qy	470 EAALELGISTIEKPGVPYHSEKMPHLVNLSDDLLAECLVYNIKPGOTRVGNVNOQTOA	529	
Db	537 -----ggslekrll--pedsd-----vamed-----acchtenhep--etvdmrtetet	576	
Qy	530 EIRLN---GSKILKEHCTF-----ENVNVVTIVPNEKAA---VMVNG	566	
Db	577 gitecnqkthsqtdhessfqplsvkdalcslnksedvsscdlvpqdvtsanvliadg	636	
Qy	567 VRIDKPTRL	575	
Db	637 v--ddpehl	643	
RESULT 14			
G40077			
ID	G40077 standard; Protein; 1460 AA.		
XX			
AC	G40077;		
XX			
DT	18-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 49678.		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
OS	Arabidopsis thaliana.		
XX			
PN	EPI033405-A2.		
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 14.4%; Score 580.5; DB 21; Length 1460;			
Best Local Similarity 29.4%; Pred. No. 2e-35;			
Matches 189; Conservative 99; Mismatches 183; Indels 171; Gaps 25;			
Qy	40	PPPGAEEKARKSKTTMDGPKAFDRSYPDKNAPYARQEDLFODLGVPLLDNAFKG	99
Db	9	pppsnplrkisaetates-----gfsds-----gkveqmqlvgaplvencisg	53
Qy	100	YNYCTFAYGQTGGKSYSMWG-----YCKEHGVIPRICODMRRINE-----	141
Db	54	fnssvfygdtgsktytmwgpangllleehlcgdgrgitprvferlfarikevmppgfvqf	113
Qy	142	----LQDKNLTCTVSVLETYNERNVRLDN-----PSTKGNL	176
Db	114	vvltnkasmklgn-sitnvghyrsfswlhyislddlqrannrptgkpekpgkghv	172
Qy	177	-----KVREHPSGTPVEDLAKLVRSFQEIENLMDEGNKARTVAATNNNETSSRS	228
Db	173	hrfslmchqdredvsgyvenleeyvknldvsgllkglgnrttgatsvntessrsh	232
Qy	229	AVFTLTITQKHDEETKMDTEKVAKISLVDLAGSERATSGATGARLKEGAENRSLSTL	288
Db	233	cvftcvvescknvødglsstfksrlnlvdlagserqkstgaagerlikeagnirsls	292
Qy	289	GRVINAALADWS-SGKQKNQLVPRDSVLTWLLKSLGNSWFTAMIAISPADINFEETL	347
Db	293	gnlilnlaelsctgkprh---lpyrdsrltflqeslgnaklamvcavspqrs--etf	347
Qy	348	STLRYAASAKRIKNHAVNE-----DPN--ARMIRELKEELAQLRSK-----	387
Db	348	stlrfaqrakaiqnkavnevmdvnlfrgvlhqlrdelqrmdgnnptnvpaysta	407
Qy	388	-----LQSGGGGGGAGSGGPPVEESYPDPPLKQ--IVSIOQPDATVKKMSKAE	436
Db	408	wnarrslnllrsftlg-----hprslphedndgdieideaaaverlcqv	453
Qy	437	IVEQUNOSEKLYRDLNQTWEELKARTEIHKREAALEELGISTEKFGFVYHSEMPHL	496
Db	454	glqslasaginhdmm-----rvksihsed-----ggsieklr1-pedsd-----	491
Qy	497	VNLSDDPLLAECVLVYKPGQTRGVNQNQDTQAEIRLN-----GSKILKEHCTF-----	545
Db	492	vamed-----acchtenep--etvdnmrtetgirenqikthsqtlidhessfqlsvkd	545
Qy	546	-----BNVDNVVTIVPNEKAA-----VMVNGVRIDKPTRL	575
Db	546	alcsalnksedvascpdlvpqdvtsanvliaadv--ddpehl	585

RESULT	15
G40076	
ID	G40076 standard; Protein; 1462 AA.
XX	
AC	G40076;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 49677.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
XX	termination sequence.
OS	Arabidopsis thaliana.
XX	
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.

```
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 14.4%; Score 580.5; DB 21; Length 1462;
Best Local Similarity 29.4%; Pred. No. 2e-35;
Matches 189; Conservative 99; Mismatches 183; Indels 171; Gaps 25;

QY 40 PPGAEKARKSGKTMDGPKAFADRSYWSFDKNAPYARQBDLFQDLGVPLLDNAFKG 99
Db 11 pppanplkrkisaetates-----gvkedmfglvgaplvencisg 55
QY 100 YNNCIFAYGOTGSGKSYSMWG-----YGREHGVIPRICODMPERINE----- 141
Db 56 fnssvfayqgsgktycmwpgangllceehlcgqgrgtprvferifarkevmpgfvqf 115
QY 142 ----LQDKNLTCTVEYSYLEIYNERVRDLN-----PSTKGNL 176
Db 116 vvltnkssmlkgn-sitnvgahysrfsdwlhylslddlqrannrptgpkpekdgkhgv 174
```



```
Qy 177 -----KVEHPSTGYPVEDAKLVYRSFOEINLMDGKNKARTVAATNNMETSSRSH 228
Db 175 hrfslmchqiredvksyvyneiteeyvknitdvsqllkglnrrtgatsvntessrsh 234
Qy 229 AVFTTLTKQWHDEETKMDTEKAKISLVDLAGSERATSTGATGARLKEGAETNRSLSTL 288
Db 235 cvftcvvestcknvadglsafktsrlnlvdlagsrqtsgaagerlkeaglnrslsqli 294
Qy 289 GVIAAALADWS-SGOKKKNQLVYRDSVLTWLLKDSLGSNSMTAMIAAISPADINFEETL 347
Db 295 gnlnlilaesqtkprh---ipyrdarlflllqeslgnaklamvcavspqrs--etf 349
Qy 348 STLRVADSARKIRKNHAVNE-----DPN--ARMIRELKEELACLRSK----- 387
Db 350 stlrfaqrakaiqnkavvnevmdvnlrgvibhgrdelgmkndgnnpnvnaysta 409
Qy 388 -----LQSSGGGGGAGSGGPGVEESYPDPDPLEKQ--IVSIOOPDATVKMSKAE 436
Db 410 wnarrslnlrsfglg-----hprslphedndgdiemeideaaaverlcqv 455
Qy 437 IVEQLNQSEKLYRDLNQTWEEKLAKTEIHKREAALEBELGISTEKGFGVGYHSEMPHL 496
Db 456 glqsslaseginhdma-----rvksihhsd-----gqsiekrll--pedsd----- 493
Qy 497 VNLSDPDLIAECLVYNIKPGQTRVGNVNDQTQAEIRLN-----GSKILKEHCTF----- 545
Db 494 vamed---acchtenhep--etvdmrtetetgirenqikthsqtdidhessfqplsvkd 547
Qy 546 -----ENVDNVVTIYPNEKAA---VMVNGVRIDKPTRL 575
Db 548 aicsslnksedvscpdlvpgdvtsanvliadgv--dopehl 587
```

Search completed: April 25, 2001, 10:12:41
Job time: 337 sec

THIS PAGE BLANK (USPTO)

Db 291 -KKRSDPFIYRDSVLTWLLKENLGGNSRTAMIAALSPADINYEETLSTLYADRTQIR 349
Qy 361 NHAVNEDPNARMIRELKEELAQRLSKLOSSGGGGGAGG----- 400
Db 350 CNAINEDPNARLIRELQEEVARLRELLMAOGLSASALEGLKTEEGSVRGALPAVSSPPA 409
Qy 401 -----SGPPEESYPPPTPLEKQVSIQQPDATVKKMSKAEIVBOLNOSEKLYRDL 451
Db 410 PVSPPSTHNGELEPSPNT--ESQI-----GPEAMERLOETEKIATL 454
Qy 452 NOTWEEKLAKEEIHKEREAALEELGISIEK--GFVPGYHSEKEMPHLVNLSDDPLLAEC 509
Db 455 NETWEEKLRKEALRMEREAALLAEMGVAVREDGGTVGVSPKKTPLHLNLEDPLMSBCL 514
Qy 510 VYNIKPGQTRVGNVQDQAIRLNGSKILKEHCTFENV-----DNVVTIVPNEKAAYV 564
Db 515 LYHIKDGTVRGVQDMD-----IKLTQFIREQHCLFRSIPQPDGVEVVVTLPECPGAEYV 570
Qy 565 NGVRIDKPTRLSRGVRIILGDFHIFRPNHPEEAREOEQSILLRHSVTNSQLGSPAPGRH 624
Db 571 NGKLVTEPLVLKSGNRIVMGKNHVRFNHPEQARLERER-----GVPPP--- 614
Qy 625 DRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREASAILGLDQKISHLTDDLDAL 684
Db 615 -----PGPPPEVDWNFQAKELLEQ--QGIDIKLE--MEKRLQDL 650
Qy 685 FDDVOKARVVRGLVEDNE---DSDS 707
Db 651 ENQYRKEKEEADLLLEQORLYADSDS 676

RESULT 2

US-08-006-676B-1
; Sequence 1, Application US/08006676B
; Patent No. 5411865
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeffrey B. Oster
; STREET: 8339 SE 57th Street
; CITY: Mercer Island
; STATE: Washington
; COUNTRY: USA
; ZIP: 98040-4906
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/006,676B
; FILING DATE: 15-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oster, Jeffrey B.
; REGISTRATION NUMBER: 32,585
; REFERENCE/DOCKET NUMBER: REED-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 232 7845
; TELEFAX: (206) 236 0205
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-006-676B-1

Query Match 20.2%; Score 814; DB 1; Length 955;
Best Local Similarity 28.1%; Pred. No. 7.4e-60;

Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;
Qy 7 IKVYVRVPFWARE--IDRGAKCIVRMENQNOTILTPP-----PGAEEKARKSKTMD 57
Db 13 VKVSVRVPLNERENNAPEGTKVYAAKAAAVVTVKVLGGSNNSGAAESMGTAARRVQD 72
Qy 58 GPKAFAPDRSWSF---DKNAPNVAROEDLFQDLGVLPLLDNAFKYNNCIFYAGTGSGK 114
Db 73 ---FQFDHVFWSVETPDACGATPATQADVFTTGYPLVGHAFDGFNSCLFAYGOTGSGK 128
Qy 115 SYSMMG-----YKHEGVIPRICODMFRINELQKKNLCTCTVEVSLEYINRVRDL 168
Db 129 TYTMGADVLSGSGNGVTPICLIEFARKASVEAOGHSRWIVELGYVEVYNERVSDLL 188
Qy 169 NPSTKG-----NLKVREHPSTGPIYVEDLAKLVVRSFQEIENLMDGEGNKARTVAATNM 220
Db 189 GKRRKGVKGGEYVVDVREHPSRGVFLGQRLVEVGSLDDVVRLEIGNGVRHTASTKM 248
Qy 221 NETSSRHAVFTLT---TQKHDEETKMDTEKVKAKISLVDLAGSERATSTGATGARLK 276
Db 249 NDRSRSHAITMLLRERTWTTSGETITRTAGKSSRNMLVDLAGSERVAOSQVEGQFK 308
Qy 277 EGAEINRSLTLGRVIAALADMS--GKOKNQLPYRDSVLTWLLKSLGGNSMTAMIAA 335
Db 309 EATHINLSLTTLGRVIDVLADMATKGAQYSPAPFRDSKLTFFILKDSLGGNSKTFMIAT 368
Qy 336 ISPADINFEETLSTLYADSARKIKNHAVNEDNARMIRELKEELAQRLSKLQSSGGG 395
Db 369 VSPSALNYEETLSTLYASRARDIVNVAQVNEPRARRIRELEQMEDMROAM----- 421
Qy 396 CGAGSGGPGVEE-----SYPPDTPLEK 417
Db 422 --AGDPAYVSELKKKLLALESEAKRAADLOALERERHNOVQVRLRLPATEAKSELES 479
Qy 418 QIVSTQQ-----PDATYKKMSKABIVBQL 441
Db 480 RAAALQEEMTATROADKMOALNRLKEEQARKERBELKEMAKKDAALSKYRRRDAEIA 539
Qy 442 NOSEKLYRDLNOTWEEKLAKEEIHKEREAALEELGISIEKGFVGPYHSEKEMPHLVNLS 501
Db 540 SEREKL-----ESTVAOLEREQEREVALDAL-----OTHORKLQEALESSE 581
Qy 502 -----DPLLAECVYNIKPGQTRVGNVQDQAIRLNGSKILKEHCTFENVNVTIV 555
Db 582 RTAAERDQLQ--LTLEQSERTQLSQVVTDRERUTR-DLQRIQYGETELAROVAALCA 638
Qy 556 PNE-----KAAVMYNGVRIDKPTRLSRGVRIILGDFHIFRPNHPEEAREOEQSILLRHS 610
Db 639 AQEMEARYHAAVFHLQTLLELATWEDALR-----ERALAERDEAAAAE-- 682
Qy 611 VTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREASAILGLD 670
Db 683 -----LDAASSTQN-----ARESACERLTSL 705
Qy 671 QKISHLTDELDALFDVQKARAVRRGLVEDNEDSDSSFPVRDKYMSNGTINFSLDT 730
Db 706 QQLRE-SEERAELASQLEATAAKSSAEQDRENTATLEQOLRES-----EARAEL 757
Qy 731 AITMPGTPRSDDDGDLFFGDKKSKQDASNV-----DVEELRQOQAOAMEEALKTKA 781
Db 758 ASQLEATAAA-----KMSAEQDRENTATLEQQLRDSEREAELASQLESTTAAK 807
RESULT 3
US-08-282-845-2
; Sequence 2, Application US/08282845
; Patent No. 5719263
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: A 230kd Antigen Present in Leishmania
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Macintosh Operating System 7.1
; SOFTWARE: Microsoft Word for Macintosh 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/006,676
; FILING DATE: JANUARY 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 5004-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-282-845-2

Query Match          20.2%; Score 814; DB 1; Length 955;
Best Local Similarity 28.1%; Pred. No. 7.4e-60;
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

QY   7 KVVVVRVPENARE--IDRGAKCIVRMGNQTITPP-----PCAEEKARKSGKTIMD 57
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   13 VKVSVRVPLNERENNAPEGTKVVAQAQAAAVTVKVLGGSNNAGAESMGSTARRVAQD 72
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   58 GPKAFAFDRSYNSF--DKNAPNAVQEDELFDGLGVPLLDNAEFGYNKNCFAYGOTGSGK 114
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   73 ----FQDFHVFVSVEPTDACCATPATQADVFTIGYFLVOHAEDFGNSCLFAYGOTGSGK 128
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   115 SYSMWG-----YGKEHGVIPIRCQDMFRINELQDNKTCTVEYSYLEIYNERNVDLL 168
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   129 TYTMWGDVSAISGEGNGVTPTCLIEIFARKASVEAQGHGRMIVELGVEVYNERVSDLL 188
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   169 NPSFKG-----NLKYREHPSTGPYVEDLAKLVSRFOEIENLMDGSKARTVAATNM 220
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   189 GKRRKGVGKGEVYDVREHSPKGVFLGGQLRLEVESLDDVVRLFEIGNVRITASTKM 248
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   221 NETSSRHAVFTTL----TQKHDEETKMDTEKVAKISILVDLAGSERATSTGATGARLK 276
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   249 NDRSSRSHATIMLLRERTTTTKSGETTIRTAGKSSRNMLVDLAGSERVAQSVEQGQFK 308
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   277 EGAEINRSLTGRIYAALADMSS-CQKKNOPLPYRDSVLTWLKLSLGSNSMTAMIAA 335
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   309 EATHINLSLTLGRVIDVLDAMTKAKAQSVAFPRDLSKLFFILKDSLGSNSTFMAT 368
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   336 TSPADINFETLSTURYADSARKNHAVVNPDNARMIREKELLAQLRSKLOSGGG 395
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   369 VSPSALNYETLSTURYASRADIVNVAQNEDPRARRIRELEEQMEDMQAM----- 421
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   396 GGAGSGGPGVEE-----SYPPDTPLEK 417
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   422 --ACGDPAYVELKKKLALLESEAQRRAADLOALERERHNQVQERLLRATPAEKSELES 479
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   418 QIVSIQQ-----PQATVKMKSAEIVVOL 441
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Qy 7 IKVVRVRFNARE--IDRGAKCIVRMENOTILTPP-----PGAEKARKSGKTTMD 57
Db 13 VKSVVRPLNERNNAPEGKTVAAKAAAVTVKVLGSSNNGAESMGTRARRAQD 72
Qy 58 GKAPAFDRSWSPF---DKNAPNTARQEDLPDGLGVPLLDNAFKYNNCIPAYQGTSGK 114
Db 73 ----FOEDHVFWSVETPDACGATPATQADVETTCYPLVQHAFDGFNSCLFAYQGTSGK 128
Qy 115 SYSMMG-----YKKEHGVIPRICQDMFRRINELQKDKNLCTVEVSYLEIYNERNVDLL 168
Db 129 TYTMGADVLSALSGEGVTPRICLEIFARKASVEAOGHSRWIVELGYVEYNERSVDLL 188
Qy 169 NPSTKG-----NLKVRHPSTGPVYEDLAKLVRSFQEIENLMDGKNKARTVAATNM 220
Db 189 GKRRKGVKGGEEYVDVREHPSRGVLEGQRLVEGSLDDVRLIEIGNVRHTASTKM 248
Qy 221 NETSSRHAVFTL---TQKHDEETKMDTEKVKLSLVDLAGSERATSGATGARLK 276
Db 249 NDRSSRSHAITMLLREERTMTTKSGETIRTAGKSSRNMLVDLAGSERVAQSQVEGQOFK 308
Qy 277 EGAENRSLTLGRVIAALADMSS-GKOKKNQOLVPRDSVLTLLKSLGNSMTAMIAA 335
Db 309 EATHINSLTLGLRVIDVLADMATKAKAQSVAFFRDSKLTFFILKDSLGNSTFMIAT 368
Qy 336 ISPADINFEETLTRYADSAKRIRKNHAVNEDPNARMIRELKEELAQRLSKLOSSGGG 395
Db 369 VPSALNVEETLTRYASRARDIVNVAQVNEPPRRIRIREEEQMEDMRQAM-----421
Qy 396 GGAGSGGPPVEESYPPDTPLEKQIVSIQOQDATVKKMSKAEIVFQLNQSEKLYDLNQTW 455
Db 422 --AGGDPAVYSE-----LKKLALLES-----EAKRAADLQALERE----REHNOVQ 463
Qy 456 EEKLAETKEIHK---REAALEE 475
Db 464 ERLLRATEAEKSELESRAAALQE 486
```

RESULT 5

US-08-428-414A-3
; Sequence 3, Application US/08428414A
; Patent No. 5912166
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TITLE OF INVENTION: LEISHMANIASIS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue

```
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,414A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kadlecck, Ann T.  
; REGISTRATION NUMBER: 39,244  
; REFERENCE/DOCKET NUMBER: 210121.407  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836 SEEDANDBERRY  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 955 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-428-414A-3
```

Query Match 20.2% Score 812.5; DB 2; Length 955;
Best Local Similarity 40.0%; Pred. No. 9.9e-60;
Matches 201; Conservative 75; Mismatches 164; Indels 63; Gaps 13;

```
Qy 7 IKVVRVRFNARE--IDRGAKCIVRMENOTILTPP-----PGAEKARKSGKTTMD 57
Db 13 VKSVVRPLNERNNAPEGKTVAAKAAAVTVKVLGSSNNGAESMGTRARRAQD 72
Qy 58 GKAPAFDRSWSPF---DKNAPNTARQEDLPDGLGVPLLDNAFKYNNCIPAYQGTSGK 114
Db 73 ----FOEDHVFWSVETPDACGATPATQADVETTCYPLVQHAFDGFNSCLFAYQGTSGK 128
Qy 115 SYSMMG-----YKKEHGVIPRICQDMFRRINELQKDKNLCTVEVSYLEIYNERNVDLL 168
Db 129 TYTMGADVLSALSGEGVTPRICLEIFARKASVEAOGHSRWIVELGYVEYNERSVDLL 188
Qy 169 NPSTKG-----NLKVRHPSTGPVYEDLAKLVRSFQEIENLMDGKNKARTVAATNM 220
Db 189 GKRRKGVKGGEEYVDVREHPSRGVLEGQRLVEGSLDDVRLIEIGNVRHTASTKM 248
Qy 221 NETSSRHAVFTL---TQKHDEETKMDTEKVKLSLVDLAGSERATSGATGARLK 276
Db 249 NDRSSRSHAITMLLREERTMTTKSGETIRTAGKSSRNMLVDLAGSERVAQSQVEGQOFK 308
Qy 277 EGAENRSLTLGRVIAALADMSS-GKOKKNQOLVPRDSVLTLLKSLGNSMTAMIAA 335
Db 309 EATHINSLTLGLRVIDVLADMATKAKAQSVAFFRDSKLTFFILKDSLGNSTFMIAT 368
Qy 336 ISPADINFEETLTRYADSAKRIRKNHAVNEDPNARMIRELKEELAQRLSKLOSSGGG 395
Db 369 VPSALNVEETLTRYASRARDIVNVAQVNEPPRRIRIREEEQMEDMRQAM-----421
Qy 396 GGAGSGGPPVEESYPPDTPLEKQIVSIQOQDATVKKMSKAEIVFQLNQSEKLYDLNQTW 455
Db 422 --AGGDPAVYSE-----LKKLALLES-----EAKRAADLQALERE----REHNOVQ 463
Qy 456 EEKLAETKEIHK---REAALEE 475
Db 464 ERLLRATEAEKSELESRAAALQE 486
```

RESULT 6
US-08-713-815A-4

; Sequence 48, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1612 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-545-860D-48

Query Match 4.8%; Score 195; DB 3; Length 1612;
Best Local Similarity 20.4%; Pred. No. 2.4e-07;
Matches 139; Conservative 90; Mismatches 241; Indels 210; Gaps 28;

QY 43 GAEEKARKSGKTIIMDGPKAFADRSYWSFKDNAPNYARQEDLFODLGVPLLDNAFKGYNN 102
DB 5 GRDEERRKLIADII-----HHWN-----ANRLDIF-EISOPTDELEFGVNR 44
QY 103 CIFYAGQTG--SGKSYMMGYGKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIY 160
DB 45 FYFDKKAAGNFATKICIRVSSTATQDVETLAE-----KFRPDMRLSSPKYSLEYVH 97
QY 161 --NERVRLNLPSTKGNLKVREHSTGYPVEDLAKLVVRSFQETENLMDENKARTVAAT 218
DB 98 VSGERRLDIDEKPLVVLWNKNDREGREV-----LKNENDAIPPKAQS 141
QY 219 NMNETSSRSHAV--FTLTLTOKWDEETKMDTEKVKAKISLVYDLAGSERATSTGATGARLK 276
DB 142 NGPEKQKEGVIONFKRTLSKKEKKRKEALQROASDKD-----DRPF 187
QY 277 EGAETNRSLTLGRVIAALADMSSGQKKQLVPYRDSVLTLWLLKDSLGLSGNSMTAMIAAI 336
DB 188 QGEDVENS-----RLAEV-----YKDM----- 205
QY 337 SPADINFEETLSTLYADSARKIKHAVNEDPNARMIRELKEAQLRSKLQSGGGGG 396
DB 206 --PETSFTRTIS-----NPEVMKRRRQKLEKRMQBFSS--D 240
QY 397 GAGSGGQVPE---BSYPDPTPLEKQIVSIQOPDAIVKKMSKAEIVEQLNQSEKLYRDL-- 451
DB 241 GRPDSGGTLRIYADSLKRNIPYKTIILLSTTDP---ADFVAEALAEKYLEKENPKDYCI 296
QY 452 -----NOTWEEKLAKTEIHKEREAALE-----ELGISI----- 480
DB 297 ARVMLPPCAQHSDEKGAKEIILDDCEPLQIFREMPWSDKGLVYFQLKRRPPDHIPKTK 355
QY 481 -----EKGFGVGYHS-----KEMPHLVNLS-----DPLLAELVYNNIKPGQT 518
DB 356 KHEGKTPKGERADGSVYGTLPPEKLPYVELSPDGSDSRDKPKL-----YRLQLSVT 410
QY 519 RVGNVNDQTAQAEIRLNGSKILKEHCTFENVNVTIVNE-KAAVWVNGVRIDKPTRLS 577
DB 411 EVGTEKLDNS-IQLFGGIQPHHCDLTNMDGVVTVTPRSMDAETVYEGQRISETTNLOS 469
QY 578 GYRIILGDFHIFRHNHPEARAERQESQLLRHSVNTNSQLSPAPGRHRTLSKAGSDAG 637
DB 470 GMKVQFGASHVFKFVDPSP-----QDHAKARSV-DGGLMVKGPRHKGPIVOETTFDLGG 522
QY 638 DSRSDSPLPFRGK---DSD 654
DB 523 DIHSGTALPTSSTRLDSD 542

RESULT 9
PCT-US94-04496-48
; Sequence 48, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04496

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: DeLuca Esq., Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJJU-1242

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 1612 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US94-04496-48

Query Match 4.8%; Score 195; DB 5; Length 1612;

Best Local Similarity 20.4%; Pred. No. 2.4e-07;

Matches 139; Conservative 90; Mismatches 241; Indels 210; Gaps 28;

QY 43 GAEKARKSGKTIIMGPKAFADRSYMSFKNAPYARQEDLFQDLGVPLLDNAFKGYN 102

Db 5 GRDEERRKLADI-----HHWN-----ANRLDLF-EISQPTDELFHGVMR 44

QY 103 CIFAYGOTG--SKSYSMWGKGVIPRICODMFRINELQKDKNLTCTVEVSYLEIY 160

Db 45 FYFOKAGNFATKICIRVSSTATTQDVETLAE-----KFRPDMRLSSPKSYLYEVH 97

QY 161 --NERVDLLNPSTGNLKVREHPSTGYPVEDLAKLVRSFQEIENLMDGNKARTVAAT 218

Db 98 VSGERRLDIDEKPLVQLWNKDDREGFV-----LKNENDAIPTPKAQS 141

QY 219 NNNETSSRSHAV--FTLTQKWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLK 276

Db 142 NGPEKQKEGVIONFRKTLSEKKEKKREKALQAQSDKD-----DRPF 187

QY 277 ECAENRSLTLGRVTAALADMSGKOKKNQLPYRDSVLTWLLKDSLGNSMTAMIAAI 336

Db 188 QGEDVENS-----RLAAEV-----YKDM-----205

QY 337 SPADINFEETLRYADSARIKNHAVNEDPNRMIRELKEELAQLRSKLSQSGGGG 396

Db 206 --PETSFTRTIS-----NPEVYMKRRQOKLEKRMQEFPS--D 240

QY 397 GAGSGGGPVE--ESYPPDTPLEKQIVSIQOPDQATVKKMAEIVQOLNOSEKLYRDL-- 451

Db 241 GRPDSGGTLRYADSLSKPNIPYKTLITLSTDDP-----ADFVAEALAEKYLEKENPKDYCI 296

QY 452 -----NOTWBEKLAETEEHKREAALE-----ELGISI-----480

Db 297 ARVMLPPCAQHSDEKGAKEIILDDDECPLQIFREWPSDKGILVFLQKRPPDHIPKTK 355

QY 481 -----EKGFVGPYHS-----KEMPHLYNLS-----DPLLAECCLVYNIKPQT 518

Db 356 KHEGKTPKGERADGVSGLTPPEKLPYLVELSPDGSDSRDKPKL-----YRLQLSVT 410

QY 519 RVGNODTOAETRLNGSKILKEHCTFENVDMVTVIPNE--KAAMVNGVRIDKPTLRLS 577

Db 411 EVGTEKLDNS--IQLFPGIQLPHCDLTNMDGVVTVTPRSMDAETTYEGQRISETTLMQ 469

QY 578 GYRIILGDFHIFRNPPEARAEQESLLRHSVTNSQLGSPACPRDRTLSKAGSDADG 637

Db 470 GKMVOFGASHVFKFVDPDS-----ODHALAKRSV--DGGLVMVKGRPHKPGIVQETTFDLGG 522

QY 638 DSRSDSPLHFRGK-----DSD 654

Db 523 DIHSGTALPTSCKSTTRLDSD 542

RESULT 10

US-09-157-420-1

; Sequence 1, Application US/09157420

; Patent No. 6180760

; GENERAL INFORMATION:

; APPLICANT: TAKAI, Yoshiaki

; APPLICANT: NAKANISHI, Hiroyuki

; APPLICANT: MANDAI, Kenji

; APPLICANT: WADA, Manabu

; APPLICANT: OBAISHI, Hiroshi

; TITLE OF INVENTION: ACTIN FILAMENT-BINDING PROTEIN "L-AFADIN"

; FILE REFERENCE: 98-1042*/LC(WMC)/653

; CURRENT APPLICATION NUMBER: US/09/157,420

; CURRENT FILING DATE: 1998-09-21

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1829

; TYPE: PRT

; ORGANISM: rat

US-09-157-420-1

Query Match 4.6%; Score 185.5; DB 4; Length 1829;

Best Local Similarity 21.1%; Pred. No. 1.8e-06;

Matches 148; Conservative 91; Mismatches 265; Indels 199; Gaps 32;

QY 43 GAEKARKSGKTIIMGPKAFADRSYMSFKNAPYARQEDLFQDLGVPLLDNAFKGYN 102

Db 5 GRDEERRKLADI-----HHWN-----ANRLDLF-EISQPTDELFHGVMR 44

QY 103 CIFAYGOTG--SKSYSMWGKGVIPRICODMFRINELQKDKNLTCTVEVSYLEIY 160

Db 45 FYFOKAGNFATKICIRVSSTATTQDVETLAE-----KFRPDMRLSSPKSYLYEVH 97

QY 161 NERVLDLLNPSTGNLKVREHPSTGYPVEDLAKLVRSFQEIENLMDGNKARTV----- 215

Db 98 -----VSGERRLDIDEKPLVQLWNKDDREGFV-----LKNENDAIPTPKAQS 141

QY 216 -----AATNNETSSRSHAV--FTLTQKWHDEETKMDTEKVAKISLVLDLAGSERAT 266

Db 133 DAIPAKKAGNSGPEKQKEGVIONFRKTLSEKKEKKREKALQAQSDKD-----DRPF 187

QY 267 STGATGARLKEGAEINRSL--STLGRVIA--ALADMSSGKOKK--NOLVPRDS----- 314

Db 188 SQGDDSENSRLAAEVYKDMPTSFTRTISNPEVYMKRRQOKLEKRMQEFPS--D 240

QY 315 VLTWLLKDSLGNN--SWTAMTAALSADINFEETLRYADSARIKNHAV--VNEDPN 370

Db 248 GTLRIYADSLKPNIPYKTLITLSTDDP-----ADFVAEALAEKYLEKENPKDYCI 296

QY 371 ARMIRE-----LKEELAQLRSKLSQSGGGGAGGGGAGGGGAGGGGAGGGGAGGGG 423

Db 306 AOHSDERGAKEIILDDDECPLQIFREWPSDKGILVFLQKRPPDHIPKTK 355

QY 424 QPDATVKKSKAEIVQOLNOSEKLYRDLNQTWEEKLAKTEEIKHKEAALEBELGISIEKG 483

Db 347 PPDYIPKMKK-----HVEGK-PLGKORADGSG 374

QY 484 FVGPHYSKEMPHLYNLS-----DPLLAECCLVYNIKPQT 518

Db 375 YGSALEPPKLPYLVELSPDGSDSRDKPKL-----YRLQLSVT 410

QY 522 NVNODTOAETRLNGSKILKEHCTFENVDMVTVIPNE--KAAMVNGVRIDKPTLRLS 577

Db 430 TEKEDDNS--IQLFPGIQLPHCDLTNMDGVVTVTPRSMDAETTYEGQRISETTLMQ 469

QY 581 IILGDFHIFRNPPEARAEQESLLRHSVTNSQLGSPACPRDRTLSKAGSDADG 637

Db 489 LQFGTSHVFKFVDPDS-----ODHALAKRSV--DGGLVMVKGRPHKPGIVQETTFDLGG 522

; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-16216-1

Query Match 3.5%; Score 143; DB 5; Length 3248;
Best Local Similarity 19.5%; Pred. No. 0.017;
Matches 141; Conservative 110; Mismatches 265; Indels 208; Gaps 30;

Qy 123 KEHGVIPRICQDMFRRINELQDK-NLTCTVEVSYL-----EIVN-----E 162
Db 2159 KEKELLVKESLSQARLSSESYKLVNSKALAEALVKEGEFALRLSSTQEEVHQLRGIE 2218
Qy 163 RVRDLLNPSTGKLV-----REHPTGPGYVEDLAKLVRS--FOEINLMDGNKA 212
Db 2219 KLRVRIEADKKQHTAEKLERERENDSLKDKVENLERELQMSQENQELVILDAENSKA 2278
Qy 213 RTVAATNNETSSRSRHAVFTLTQKHDEE--TKMDTEKVAKISLVDLAGSERATSTGA 270
Db 2279 EVETLKTOIEMARSLKIFELDLVTLRSEKENLTQIQKQOGSLED-----KLLSFK 2333
Qy 271 TGARLKEGAEI-----NRSLSLGRVIAALADMSGKQKKNQLVPYRDSVLT 317
Db 2334 SLLEKEQAEIQTKESKTAVEMQLNQKELNEAVALC-----GDQE----- 2376
Qy 318 WLLKDSLGNSMTAMIAISPADINFETLSLRYADSASAKRIKHAVVNDPNARMREL 377
Db 2377 -----IMKATEQSLOP-PIEEHQL-----NRSIEKLRLARLEADEKKKOLCVLOQL 2420
Qy 378 KEELAQRLSKLQSSGGGGAGSGGPGVEESYPPDTPLEKQIVSIQPDATVK-KMSKAE 436
Db 2421 KE--SEHADLLK-----GRVE-----NLERELEIARTNQEHAALEAENSKGE 2461
Qy 437 I-----VEQLNQSEK-LYRDLNQTWEEKLAKEEIIHKEREAALEELGISIEKGFVGPYH 489
Db 2462 VETLAKIEGTQSLRGLELDVVTIRSEKENLTNQLQKQE-RISELEI-INSSFNILQ 2519
Qy 490 SKEMPHLVNLSDDPLLAECVLNPKGTGVGNVQDTQAEIRLNGSKILKECTF--EN 547
Db 2520 EKEQEKVQMKESSTAMEMLQTQLKELNERNVAALHNDQEA-----CKAKQON 2566
Qy 548 VDNVTVIPNKAAYVNGVRIDKPTRLRSGYRIILGDFHIFRNHPPEARAERQESLL 607
Db 2567 LSSQVECLELEKALLOQ---LDE---AKNNYIVL-----QSSV 2599
Qy 608 RHSVNTSOLGSPAPGRHRTLSKAGSDGDSPLPFRGKDSDFYAREREAAIL 667
Db 2600 KGLIOEVDGKQLEKDEISRLKNQIQOEOGLVSKLSQVEGHEQLWKQONLEIRLNTV 2659
Qy 668 GLQOKISHL-----TDELDALFDVQKARAVRGLVEDNEDSDSPPVVRDKYMSNG 721
Db 2660 ELEOKIQLVQSKNASLQDTLEVL-----QSSV----- 2686
Qy 722 TIDNFSLDATATMPGTPRSDDDGALFFGDKSKODASNVDE---ELRQOQAOEERAL 777
Db 2687 --KNLENELELTQM-----DKMSFEKVNKWTAKETELQREHMAQKTAELQEL 2735
Qy 778 KTAQ 781
Db 2736 SGEK 2739

RESULT 13
US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa

; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-328-254-6

Query Match 3.5%; Score 142; DB 1; Length 2482;
Best Local Similarity 19.6%; Pred. No. 0.014;
Matches 142; Conservative 108; Mismatches 266; Indels 208; Gaps 30;
Qy 123 KEHGVIPRICQDMFRRINELQDK-NLTCTVEVSYL-----EIVN-----E 162
Db 1431 KEKELLVKESLSQARLSSESYKLVNSKALAEALVKEGEFALRLSSTQEEVHQLRGIE 1490
Qy 163 RVRDLLNPSTGKLV-----REHPTGPGYVEDLAKLVRS--FOEINLMDGNKA 212
Db 1491 KLRVRIEADKKQHTAEKLERERENDSLKDKVENLERELQMSQENQELVILDAENSKA 1550
Qy 213 RTVAATNNETSSRSRHAVFTLTQKHDEE--TKMDTEKVAKISLVDLAGSERATSTGA 270
Db 1551 EVETLKTOIEMARSLKIFELDLVTLRSEKENLTQIQKQOGSLED-----KLLSFK 1605
Qy 271 TGARLKEGAEI-----NRSLSLGRVIAALADMSGKQKKNQLVPYRDSVLT 317
Db 1606 SLLEKEQAEIQTKESKTAVEMQLNQKELNEAVALC-----GDQE----- 1648
Qy 318 WLLKDSLGNSMTAMIAISPADINFETLSLRYADSASAKRIKHAVVNDPNARMREL 377
Db 1649 -----IMKATEQSLOP-PIEEHQL-----NRSIEKLRLARLEADEKKKOLCVLOQL 1692
Qy 378 KEELAQRLSKLQSSGGGGAGSGGPGVEESYPPDTPLEKQIVSIQPDATVK-KMSKAE 436
Db 1693 KE--SEHADLLK-----GRVE-----NLERELEIARTNQEHAALEAENSKGE 1733
Qy 437 I-----VEQLNQSEK-LYRDLNQTWEEKLAKEEIIHKEREAALEELGISIEKGFVGPYH 489
Db 1734 VETLAKIEGTQSLRGLELDVVTIRSEKENLTNQLQKQE-RISELEI-INSSFNILQ 1791
Qy 490 SKEMPHLVNLSDDPLLAECVLNPKGTGVGNVQDTQAEIRLNGSKILKECTF--EN 547
Db 1792 EKEQEKVQMKESSTAMEMLQTQLKELNERNVAALHNDQEA-----CKAKQON 1838

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:09:24 ; Search time 46.78 Seconds
(without alignments)
1151.748 Million cell updates/sec

Title: US-09-235-416-1
Perfect score: 4030
Sequence: 1 MSGGNIKVVVRVRFNARE.....ELRQQQAQMEALKTAQEF 784

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	41.5	1150	1 A55289	kinesin-like prote
2	1663.5	41.3	1695	2 A56921	kinesin family pro
3	1593.5	39.5	1584	1 JN0114	kinesin-related pr
4	1590.5	39.5	1584	2 T15822	kinesin-like prote
5	1396.5	34.7	1921	2 T13827	kinesin-73 - fruit
6	1284.5	31.9	1576	2 T29237	hypothetical prote
7	1147	28.5	928	2 T16759	hypothetical prote
8	1035	25.7	1121	2 T13796	kinesin-like prote
9	1034	25.7	1121	2 T13796	kinesin-like prote
10	860	21.3	699	1 S38982	kinesin-related pr
11	844.5	21.0	747	1 A57107	kinesin-related pr
12	828.5	20.6	742	1 S58691	kinesin-related pr
13	825.5	20.5	786	2 A53939	kinesin homolog KH
14	814	20.2	955	2 A47334	Lckin kinesin-rela
15	800	19.9	701	1 B44259	kinesin-related pr
16	782.5	19.4	1231	2 A34803	microtubule-associ
17	775	19.2	1229	2 T48959	kinesin-like prote
18	772	19.2	928	2 T10164	kinesin heavy chai
19	765.5	19.0	793	2 JC5831	kinesin-related pr
20	755.5	18.7	1225	2 A56514	chromokinesin - ch
21	746.5	18.5	784	1 A55236	kinesin-related pr
22	744	18.5	1226	2 T51617	kinesin-like prote
23	743	18.4	935	2 T51930	kinesin [imported]
24	722	17.9	929	2 T51932	kinesin [imported]
25	708	17.6	672	2 S54351	kinesin osm-3 - Ca
26	701	17.2	1130	2 T21134	hypothetical prote
27	691.5	17.2	963	1 A41919	kinesin heavy chai
28	689.5	17.1	294	2 S38983	kinesin-related pr
29	687.5	17.1	1031	1 A38713	kinesin heavy chai

30 687 17.0 881 2 184737 kinesin heavy chai
31 684.5 17.0 1032 2 138510 neuronal kinesin h
32 684 17.0 967 1 A35075 kinesin heavy chai
33 681 16.9 968 2 T51933 kinesin motor prot
34 678 16.8 1075 2 T45570 kinesin-like prote
35 677 16.8 1027 2 S37711 kinesin heavy chai
36 671.5 16.7 843 2 S44868 kinesin heavy chai
37 671.5 16.7 1056 1 G02157 kinesin-like spind
38 668 16.6 2663 1 S28261 centromere protein
39 665 16.5 670 2 T29898 kinesin protein OS
40 663 16.5 332 2 C48835 kinesin-like prote
41 662.5 16.4 2954 2 T14156 kinesin-related pr
42 661.5 16.4 1388 2 T30335 KLP2 protein - Afr
43 649.5 16.1 975 1 A31497 kinesin heavy chai
44 635.5 15.8 1083 2 T26844 hypothetical prote
45 633 15.7 1184 1 A34795 kinesin-related pr

ALIGNMENTS

RESULT 1
A55289
kinesin-like protein KIF1B - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A55289
R:Nangaku, M.; Sato-Yoshitake, R.; Okada, Y.; Noda, Y.; Takemura, R.; Yamazaki, H.; H
Cell 79, 1209-1220, 1994
A:Title: KIF1B, a novel microtubule plus end-directed monomeric motor protein for tra
A:Reference number: A55289; MUID:95094296
A:Accession: A55289
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1150 <NAN>
A:Cross-references: GB:DI7577; NID:9407338; PIDN:BAA04503.1; PID:g407339
C:Superfamily: kinesin-related protein KIF1B; kinesin motor domain homology
C:Keywords: P-loop
F:6-354/Domain: Kinesin motor domain homology <KMOF>
F:97-104/Region: nucleotide-binding motif A (P-loop)

Query Match	41.5%	Score 1674;	DB 1;	Length 1150;
Best Local Similarity	43.4%;	Pred. No. 6.5e-83;		
Matches 380;	Conservative 137;	Mismatches 237;	Indels 122;	Gaps 19;
QY 4	GNIKVVVRVVRPFPNAREIDRGAKCIVRMENQOTILTPPPGAEEKARKSKCTIMDGPKAFA 63			:
Db 3	CASVKVAVRVVRPFPNSRETSKESKCIIOQGNSTSIINPKPKKE-----APKSFS 51			:
QY 64	FDRSYWSEF-DKNAPNVARQEDLPDGLVPLLDNAPKGYNNCFAYGOTGSGKSYSMGYG 122			:
Db 52	FYISYSHTSPEPDCPCFASQNRVYNDIGKEMLLHAFEGYNNVCFAYGQTGAGKSYTMKGQ 111			:
QY 123	KEHG--VIPRICQDMFRINELQKRLNLTCTVEVSYLEIYNERVRLDLPNPSTKGNLKVRE 180			:
Db 112	BESQAVIIPOLCEELFEKIND-NCNEEMSYSEVSEYMEIYCERVRLDLPNPKNGNLVRE 170			:
QY 181	HPSTGPVVEDAKLVVRVRFQIEINLMDENGNKARTVAATNNMETSSRSRSHAVFTLTOKWH 240			:
Db 171	HPLLPYVEDLSKLAVTSTYTDIADLDAGNKARTVAATNNMETSSRSRSHAVFTIVTQKKQ 230			:
QY 241	DEETKMTKRVAKISLVDLAGSRATSGATCARLKEAGINSLTGLGRVIALADNWS 300			:
Db 231	DPEINLSTKRVKISLVDLAGSRADSTGAKTRLKEGANINKSLTLGKVISALAEVSK 290			:
QY 301	GQKQKNOLVPRDSVLTWLLKDSLGGSNTAMIAAISPADINFEETLSTLRVYDSAKRIK 360			:
Db 291	-KKKKTDFIPYRDSVLTWLLRENIGGNSRTAMVAALSPADINDETSLTLRYADRAKQIK 349			:
QY 361	NHAVNPDNARMRELKEELAQRLKSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 418			:
Db 350	CNAVINEDPNKLVRELKEEVYRLKDLLRAQGLGDIIDTSMGSLT--SSPSSCSLSNQVG 407			:

A:Molecule type: mRNA
 A:Residues: 1-1921 <LH>
 A:Cross-references: EMBL:U81788; NID:g1906595; PID:g1906596; PIDN:AAB50404.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0019968

```

Query Match      34.7%; Score 1396.5; DB 2; Length 1921;
Best Local Similarity 48.8%; Pred. No. 1.4e-67;
Matches 305; Conservative 88; Mismatches 171; Indels 61; Gaps 8;

Qy 7 IKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKIMDGPKAFADR 66
Db 6 IKVAVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKIMDGPKAFADR 66
Qy 67 SYWSPDKNAPYARQEDLFQDLGVLPLDAPKGYNNCFAYGQTGSGKSYMMVGKHEHG 126
Db 59 CFYSLNPEDENFASQETVDFDCVGRGILDNAPQGVNACIFAYGQTGSGKSYMMVGKHEHG 118
Qy 127 VIPRICQDMFRRLNLOKDKNLTCTVEVSYLEYIYNERVLDLNPSTKGN-LKVRHPSTG 185
Db 119 IIPRCDGLFSAIAN-KTPELMYKVEVSYMEIYNEKRVHDLDPKPKOSLKVREHNYMG 177
Qy 186 PYVEDLAKLVRSFOEINLMDDEGNKARTVAATNNNETSSRSNAVFTLTQKWHDEETK 245
Db 178 PYVDGLSOLAVTSYQIDNLTGEGNKSRTVAATNNNAESSRSNAVFTLTQKWHDEETK 237
Qy 246 MDTEKVAKISLDLAGSERATSTGATGARLKEGAEINSLTGLRVIAALADMSSGKOK 305
Db 238 VSGEKVSRMSLDLAGSERAVKTGAVGRLEKESNINKSLATLGLVLSKLADQSNKRGSG 297
Qy 306 N-QLVYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLTLRYADSAKRKNHAY 364
Db 298 NDKFPYRDSVLTLLKDSLGNSMTVMVATISPADNFEETLTLRYADSAKRKNHAY 357
Qy 365 VNEDPNARIMRELKELAQRLKQSSGGGGGAGGGGPPVEESYPPDTPLEKQIVSIQ 424
Db 358 VNEDPNARIIRLREHEVETLSMLKHA-----TSPV----- 389
Qy 425 PDATVKMSKAETVQLNOSKLYRDLNQTWEELAKTEIHKEREAALEELGISIEKGF 484
Db 390 -----GDVQDKLAESENLMKQISQWEEKLVKTERIQNERQQALEKMGISVQAS- 438
Qy 485 VGPYHSEMPHLVNLSDPLLAECIVYKPGQTRVGNQDTQAEIRLNGSKILKEHCT 544
Db 439 -GIKVKKNKYLVLNADPSLNELLVYIKDRTLIGGRTISGQQPDIDQLSLGIGQPEHC 497
Qy 545 FENVNVVTIYPNEKAAVWNGVRIDKPTRLSRGVRIILGDFHFRFNHPE----- 595
Db 498 ITIEDSGLYMEPVQGCARFVNGSAVEKTPLONGDRILNGHHFRVNSPKSNNTSMCAS 557
Qy 596 --EAAERQESLRLHSTVNSOLGS 618
Db 558 EPQTPAQLIDYNFARDEIMQNELSN 582

```

RESULT 6
 T29237
 hypothetical protein F56E3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T29237
 R:Gatung, S.; Wu, X.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F56E3.
 A:Reference number: 220592
 A:Accession: T29237
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1576 <GAT>
 A:Cross-references: EMBL:U41536; PIDN:AAB52613.1; GSPDB:GN00028; CESP:F56E3.3
 A:Experimental source: strain Bristol N2; clone F56E3
 C:Genetics:

A:Gene: CESP:F56E3.3
 A:Map position: X
 A:Introns: 23/1; 47/2; 125/3; 162/2; 191/3; 274/2; 296/3; 532/1; 651/3; 723/2;

```

Query Match      31.9%; Score 1284.5; DB 2; Length 1576;
Best Local Similarity 38.2%; Pred. No. 1.2e-61;
Matches 323; Conservative 120; Mismatches 217; Indels 185; Gaps 23;

Qy 7 IKVVVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKIMDGPKAFADR 66
Db 10 VKVAVRPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKIMDGPKAFADR 66
Qy 67 SYWSPDKNAPYARQEDLFQDLGVLPLDAPKGYNNCFAYGQTGSGKSYMMVGKHEHG 126
Db 58 SFCSTDPHSYDFASQETVDFDCVGRGILDNAPQGVNACIFAYGQTGSGKSYMMVGKHEHG 117
Qy 127 VIPRICQDMFRRLNLOKDKNLTCTVEVSYLEYIYNERVLDLNPSTKGN-LKVRHPSTG 185
Db 118 IIPVCNDIFTRIQE-TSNSSLSFKVEVSYMEIYNERVLDLDPKPKSSKALKVREHKLIG 176
Qy 186 PYVEDLAKLVRSFOEINLMDDEGNKARTVAATNNNETSSRSNAVFTLTQKWHDEETK 245
Db 177 PMVDGLSILAVNSPQISNLLDEGNKRSRTVAATNNNAESSRSNAVFTLTQKWHDEETK 236
Qy 246 MDTEKVAKISLDLAGSERATSTGATGARLKEGAEIN----- 282
Db 237 FSGEKVAKISLDLAGSERATSTGATGARLKEGAEIN----- 296
Qy 283 -----RSLSLTGLRVIAALADMSSGKOKNOLVPYRDSVLTLLKDSLGNSMTA 331
Db 297 VYVFNQKLDPRSLTLLGWISALAEKNS---KKDKFPIPYRDSVLTLLKDSLGNSRTV 353
Qy 332 MIAAISPADINFEETLTLRYADSAKRKNHAYVNEPNARMRELKELAQRLKQSS 391
Db 354 MIATLSPADINFEETLTLRYADSAKRKNHAYVNEPNARMRELKELAQRLKQSS 407
Qy 392 GGGGGAGGGGPPVEESYPPDTPLEKQIVSIQDQATVKMSKAETVQLNOSKLYRDL 451
Db 408 -----MQITQTKKEAETE-----ELRERLAESERLVAQM 437
Qy 452 NOTWEELAKTEIHKEREAALEELGISIEKGFVGPYHSEMPHLVNLSDPLLAECIVY 511
Db 438 NKSWEELKEDTINKERQKDLTEIGISIESS--GIKVEKDRFYLVNNAEDSLNELLY 495
Qy 512 NIKPGQTRVGN-----VNQDT-----QAETIRLNGSKILKEHC--TFEN 547
Db 496 YIN-GSAIIGNSEELTSRDSLSMTCSDDSRDODKERTSIVLRGLGIMRRHAKMTVEE 554
Qy 548 VDNVVT--IYP-NEKAAVWNGVRIDKPTRLSRGVRIILGDFHFRFNHPEARERQEQ 604
Db 555 YGGRLELVAPMSSECRICVNGKQITERTLLRNGRLVLGMNHFVKVNCP---KYMDMEQ 611
Qy 605 SLLRHSVTNSQLGSPAPGRHRTLSKAGSDADGDSRSDPLPHFRGKDSDFYARREAS 664
Db 612 SIMEDSTMFY-----NDAWHEVNDANPI-----SS 637
Qy 665 AILGLDQKISHLTDDLEDAFDVQKARAVRGLVEDNEDSDS--QSSFPVVRDKYMSNGT 722
Db 638 AV---DQWESVT-----LKHQEDKKAALQQYEAFAEKYIQSIT 673
Qy 723 IDNFSLDTAITMPG-----TPRSDDDGDALFFGDKKSKQDASNVVDVEELRQQAQMEALK 778
Db 674 AGGPTTSPMT-PGFCLPTITPTGLPFPFPANPKQSVKSKFFYWAQRKEEMPAESIK 732
Qy 779 TAKQE 783
Db 733 RLKAD 737

```

RESULT 7
 T16759
 hypothetical protein R144.1 - Caenorhabditis elegans

RESULT 8
T13750
linesin-11
C:Species: 1
C:Date: 13
C:Accession: 1
R:Okura, J.
J. Cell Sci
A>Title: M
A:Reference
A:Accession
A>Status: 1
A:Molecule
A:Residues
A:Cross-re
C:Genetics
A:Cross-re
A:Map posi
C:Superfam

[illegible]

```
Db 750 DHPVDFQLAHOEILKOKBOQQLRSELEAEKRAALTKIEQRAOCHARDFERLQCLELEQFK 809
Qy 650 GK-DSDFYARREASAILGLDOKISHLTDDDELALFDVQKA---RAVRGLVEDNEDS 705
Db 810 YKCNSEMLETERQA---LALAQOQEHPTLRHEDAVSTPAQKSTILEDIORIMLNPSEES 865
Qy 706 DSQSFPPVRD-----KYSNGTIDNFSL 728
Db 866 LHKQLMKVKEATQRCQLDLPLEFROTQTPDEFGL 900

RESULT 9
Tl3796
Kinesin-related protein 38B - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: Tl3796
R:Molina, I.; Baars, S.; Hales, K.; Fuller, M.T.; Ripoll, P.
J. Cell Biol. 139, 1361-1371, 1997
A:Title: A chromatin associated kinesin-related protein required for normal mitotic chr
A:Reference number: 217759; MUID:98060833
A:Accession: Tl3796
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1121 <MOL>
A:Cross-references: EMBL:Y15247; NID:g2578010; PIDN:CAA75531.1; PID:g2578011
C:Genetics:
A:Gene: tlo
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match 25.7%; Score 1034; DB 2; Length 1121;
Best Local Similarity 33.8%; Pred. No. 2.7e-48;
Matches 274; Conservative 130; Mismatches 290; Indels 116; Gaps 21;

Qy 6 NIKVVVRVPPNAREIDRG-AKCIVRMEGNOTILTPPPGAEEKARKSKTTMDGPKAFAP 64
Db 120 NWIVAVRVRPLNALECTRGQVTVVVOVHGNSNELIVQAGSSADA-SAGVT-----HFFSY 173
Qy 65 DRYSWFDKPNAPYARQEDLFDQGLVPLLDNAFGKNCIFAYGQTGSGKSYMMGY--- 121
Db 174 DOVYSCDPERKNPACQAKVEGTARPLIDTAFAEGYNACLPAYGQTGSGKSYMMGIAL 233
Qy 122 -----GKEH---GVIPRICQDFRRINELQKDKNLCTCTVEVSLEYIYNERVDLLN-- 169
Db 234 DDAALDGGPPHYEAGIIPRFCHELFRTEAVKSOQQLQVEVEVSFEIYNEKIHDLSVQ 293
Qy 170 -----PSTGNLKVREHPSTGPVYVEDLAKLVRSFQEIENLMDGNGKAR 213
Db 294 HAAATGEGSTPIQQOQQOORPALKVREHPIFGPYVVDLSAHSVDSYSAIRNLWLVAGNSQR 353
Qy 214 TVAATNMNETSRSRAVF--TITLTQKHDEETKMDTE-----KVAKISLVLDLAG 261
Db 354 ATASTAMDKSSRSHSIFNIVNLTDLSDDGLSDTDSSTASSLRQTRSKISLVLDLAG 413
Qy 262 SERATSTGATCARLKEGAENRSLTGLRVTAALADMS-----GKQKNQLVPYRDS 314
Db 414 SERISVSGNGERIREGYSINKSLTLTKVIAALADSKAANGPLGSGTPTSTFPYRES 473
Qy 315 VLTLLKDSLGNSTMATAISPADINFEETLSPLRYADSAKRIKHAHVNVNEOPNAMI 374
Db 474 VLTLLRENLGNSKTVMLATISPASIHADETLATFLRYACKARSIVNRKVNESPHDKII 533
Qy 375 RELKEELAQRLSKLQSSGGGGGAGGGGPPVEESYPPDTPLEKQTVTSQQPDATVKKMSK 434
Db 534 RDLRAEVDRLKSLRNEYERQLSGNSNPNVPRKIIET-----SVDETEVEALRQOL 586
Qy 435 AEIVPOLNQSEKLYRDLNQTWEEKLAKTEETHKEAREAALEELGISIEKFGVPHSKEMP 494
Db 587 AERERELSRQK-----SMWEKLKEADQKRSRLYKRRGLALEL-----TAEQKQA 634
Qy 495 HLVNLSDDPLLAELCLVYNIKPGQTRVAG-----NVNQDTQAEIRLNGSKILKEHCTFENV- 548
```

```
Db 635 CLVNLTA DPILLSGTLFYLLPQGLVRIGRRLPGGSSSQDPDIVLDGPLVALQHCIEHER 694
Qy 549 -DNVVTVIYVNPKEAAVMVNGVRIDKPTRLRSYRIILGFHIFRFNHP---BEARAERQEQ 604
Db 695 GKKLYVIPGSEDFEYVNGELLKDRRQLFHGDRLVIGGSHYFRISNPFCSQRGKADHPVD 754
Qy 605 SILRH-----SVTNSQLGSPAGRHDRITLSKAGSDAGDSRSDS-----PLPHPRGK-DS 653
Db 755 POLAHOEILQKOEQQLRSELEAEKRAALTKIEQRAOCHARDFERLQCLELEEQFYKCN 814
Qy 654 DWFYARREASAILGLDOKISHLTDDDELALFDVQKA---RAVRGLVEDNEDSDSSQ 710
Db 815 EMLETERQA---LALAQOQEHPTLRHEDAVSTPAQKSTILEDIORIMLNPSEESLHKTQ 870
Qy 711 FPPVRD-----KYSNGTIDNFSL 728
Db 871 LMVKEATQRCQLDLPLEFROTQTPDEFGL 900

RESULT 10
S38982
Kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)
N:Alternate names: kinesin-2 chain A; KRP (85/95) 85K chain
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S38982; S72551
R:Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.
Nature 366, 268-270, 1993
A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.
A:Reference number: S38982; MUID:94050179
A:Accession: S38982
A:Molecule type: mRNA
A:Residues: 1-699 <COLL>
A:Cross-references: EMBL:L16993; NID:g295245; PIDN:AAAL6098.1; PID:g295246
A:Accession: S72551
A:Molecule type: protein
A:Residues: 2-5,'X',7-11;59-64;125-132;222-226,'X',228-230 <COL2>
C:Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5
C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C:Keywords: ATP; heterotrimer; microtubule binding; P-loop
F:11-348/Domain: kinesin motor domain homology <KMOT>
F:97-104/Region: nucleotide-binding motif A (P-loop)
F:103/Binding site: ATP (Lys) #status predicted

Query Match 21.3%; Score 860; DB 1; Length 699;
Best Local Similarity 40.0%; Pred. No. 3.7e-39;
Matches 205; Conservative 88; Mismatches 155; Indels 64; Gaps 15;

Qy 2 SGGGNIKVVVRVPPNAREIDRGAKCIVRM-EGNOTILTPPPGAEEKARKSKTTMDGPK 60
Db 6 SGNDNVVRVRCRPLNSKETGGGFKSVVKMDMRGTQVTPNPA-----PSG-----EPPK 56
Qy 61 AFAPDRSWSFDKPNAPYARQEDLFDQGLVPLLDNAFGKNCIFAYGQTGSGKSYMMWG 120
Db 57 SFTFDTFV-----APG-AKQTDVYNQTPARPIDVAIIIEGYNGTIFAYGQTGKTFTTMEG 109
Qy 121 YGKE---HGVTIPRICQDMFRRINELQKDKNLCTCTVEVSLEYIYNERVDLLNPSFKGNL 177
Db 110 VRSQPELGILPNSEFAHIFGHI--AKEQENVFLVRSVLEYIYNEEVKDLGKDOOHRLE 167
Qy 178 VREHPSTGPVYVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSRSRAVFTLTITQ 237
Db 168 VKERPDPGVYVVKDLISAFVYVNNADDMDRIMTLGNKRSVGATNMNESSSSRSHAFITTLER 227
Qy 238 KWHDEETKMDTE---KVAKISLVLDAGSRATSTGATCARLKEGAENRSLTGLRVTA 294
Db 228 ----SDMGLDKEQHVVRGKLHMVDLAGSEROTKTGATGRLKEATKINLSLSTLGNVISS 283
Qy 295 LADASSGKQKNQLVPYRDSVLTLLKDSLGNSTMATAISPADINFEETLSRLYAD 354
Db 284 LVD-----GKSTHIPYRNSKLTRLQLDLSLGNKTVMCANIGPAEYNYDETISTLRVAN 337
```

Db 336 RAKWIKWKPVEDPDKALLRBFQBEIARLKAQLEKRGKRRKRREKRGSGSGGGEE 395

QY 405 VE---ESYPDPDPLEQVTSIQOPATVYKKMSKAEIVE--QLNSEKLYRLDNLQWTEKL 459

Db 396 BEEGEBEGEDGDDKDDYQWEQOEKLETEKRA---IVEDHSLVAEEKM-----RLIKE 447

QY 460 AKTEIHKREAALEELGISIE---KGFVGPY---HSEMPHLVNLSDLPDLAECLEVY 511

Db 448 KKMEDLPREKDAE-EMLGAKIKAMESKLLVGKNIYDHTNEQOKILSQKQRIEAG----- 502

QY 512 NIKPGQTVGVNGVNDQTAETRLNGSKILKEHCTFENVDVNTVTPNKAAVWVNGVRI DK 571

Db 503 -----KRREREQQ--QMESRDEFTLELKE--TVTSLOQVEDI-----KTKK 540

QY 572 PTRLRSYRIILGDPIFRFHNHPEARAEQES-----LLRHVSNTSOLG----- 617

Db 541 LKLFSLKLAQVKAETHDLQEEHKEHQELETQNELTRLKHLHLIENPILEKNKIM 600

QY 618 -----SPAPGRHDFTLSKAGSDAGDSRDS-----PLPFRGKDS 653

Db 601 NRSFFDDEHHKHLPTIRLENGOMMKRPVSAVGYKRPILSOHARMSWMIRPEPYRAEN- 659

QY 654 DMFYARREAAAILGLQKTIHSHTDDELDAFLDDQKARAVRGLVEDNDDSDQSFF 711

Db 660 -----IMLELDMPSTRTRDYEGPAISPQVQ--AALDAALQDEDEIQVDASSF 705

RESULT 12

S58691

N;kinesin-related protein KRP95 - sea urchin (Strongylocentrotus droebachiensis)

N;Alternate names: kinesin-2 chain B; KRP (85/95) 95k chain

C;Species: Strongylocentrotus droebachiensis

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: S58691

R;Rashid, D.J.; Wedaman, K.P.; Scholey, J.M.

J. Mol. Biol. 232, 157-162, 1995

A;Title: Heterodimerization of the two motor subunits of the heterotrimeric kinesin.

A;Reference number: S58691; MUID:95404610

A;Accession: S58691

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A;Molecule type: mRNA

A;Residues: 1-742 <RAS>

C;Complex: heterotrimer of a 115k chain and two kinesin-related chains of 85k (PIR:S3

C;Superfamily: kinesin-related protein KIF3; kinesin motor domain homology

C;Keywords: ATP; heterotrimer; microtubule binding; P-loop

F;9-345/Domain: kinesin motor domain homology <RNOT>

F;95-102/Region: nucleotide-binding motif A (P-loop)

F;101/Binding site: ATP (Lys) #status predicted

Query Match 20.6%; Score 828.5; DB 1; Length 742;

Best Local Similarity 31.5%; Pred. No 2e-37;

Matches 250; Conservative 122; Mismatches 286; Indels 135; Gaps 25;

QY 7 IKVVVRFPNFAREIDRGAKCIVRMEGNQTI--LTPPGABEAKRSGKTI MDGPKAF 64

Db 9 KVVVVRCRPMNSKETSGCHKRIVEMDNKRGIVETNPKGPPCEPKS-----FTF 58

QY 65 DRSY-RSFDKNAPYARQEDLFDGLVPLDNLAFKYNNCIFAYGOTGSGKSYSMVGX 123

Db 59 DTVYDWN-----SKQIDLDEYFTRSLVESVLOGFGNGTIFAGTGTGKTFTEGVRS 110

QY 124 E---HGVIPRICDMFRINELQKDNLTCTVEVSYLEYINRERVRLDNLNPKSTGNLKVRE 180

Db 111 NPELRGVIPNSPEHFTHIARTQNOQFI--VRASYLEYIQYEIRDLADQKRLDLKE 167

QY 181 HPSTGYPVEDLAKLVRSFOETENLMDNGNKARTVAATNMNETSSRSRAVFTLTLTQKWH 240

Db 168 RPTGVYVADLSSFTVSKVEIHTVTVGNKNRSGVTNMNHSRSHAFIITI----E 223

QY 241 DEETKMDTE---KVAKISLVDLAGSERATSTGATGARLKEAGAEINSLSTLGRVTAALD 297

Db 224 CSELGVDGNHNIRYKGLNLYDLAGSERQAKTGATGRKLEATKINLSLALGNVLSALV 283

```
QY 298 MSGGKQKQNLVYPYRDSVLTWLLKDSLGGNSMTAIAAISPADINFEETLSTLRYSADSAK 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 -----GKSSHIPYRDKSLRLLQDSLGGNAKTVYVAMNMGPSYFDETTITLYANRAK 337

QY 358 RIKNHVYVNEPDNARMIRELKEELAAQLRSKLQSSG-GGGGAGGSGGPGVEESYPDPDTPL 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 NIKNPKINEDPKDALLREFQEEISRLKQALDKKGPSDGRKKKKRPGEGGGDDIDE 397

QY 417 KQIVSIQDPATVYKMKSAEIVEQLNSEKLYRDLNLTWEEKLAKTEEI-----H 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 TEEGEDDEEEMTKESQKLEE---EKEKIMANQSMIAEEKQLLSEVOKRQGEIRKEH 454

QY 467 KERAALAEELGISTEKGFGPY----HSEMPHLVNLSDOPLLAECVLVYNIKPGQTRGVN 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 455 QOKELSGKIKAMESKLLVGKSVIHTNEQORKIE-EQRLLLAEE-----EKN 501

QY 523 VNQDTQAEIRLNGSKILKEHCTFENVDNVITVINEKAAVMVNGVRIDKPTLRSGYRII 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 502 RERDERKLEQDDKTVEIEGTFSSLOQVE-----VTKKLLKLFKQLQSY 548

QY 583 LGDPHIFRPHPEAR-AERQESLLRHVSVNSOLGSPAGRHDR--LSKAGSDADGDS 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 549 KSDIQDLQDEHARERQELEQTQNELRELKLVKVIADNFIPEERTKITTRAVFDEETEE 608

QY 640 RSDPLPHFGKD-----SDWFYAR-----REASAILGLDQKIS 674
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 609 WLLTFLAKAEGPSOMAKRPVSAGVGNRRPIAD--YARMAOQMGPNRYKAENILSLDMP 666

QY 675 HLT--DDELDAFDVQKARAVRGLVEDNEDSDQSFPV-----RDKYSNGTID 724
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 667 NXXTRDYEGSVAPRVQ---AALDAALQDEDDLLEQVPEVFKAKTKLKDKVRSK---- 719

QY 725 NFSLDTAITMPGT 737
   : : : :
Db 720 ---HKAVAKPGS 728

RESULT 13
A53939
kinesin homolog KHP1 - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 10-Dec-1999
C:Accession: A53939
R:Walther, Z.; Vashishtha, M.; Hall, J.L.
J. Cell Biol. 126, 175-188, 1994
A:Title: The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous protein.
A:Reference number: A53939; MUID:94299638
A:Accession: A53939
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-786 <NAL>
A:Cross-references: EMBL:L33697; NID:9497696; PIDN:AAA21738.1; PID:9497697
A:Note: authors translated the codon AAC for residue 753 as Asp
C:Genetics:
A:Gene: FLA10
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; coiled coil; p-loop
F:11-359/Domain: kinesin motor domain homology <KMOT>
F:97-104/Region: nucleotide-binding motif A (P-loop)

Query Match 20.5%; Score 825.5; DB 2; Length 786;
Best Local Similarity 39.3%; Pred. No. 3.2e-37;
Matches 204; Conservative 80; Mismatches 168; Indels 67; Gaps 14;

QY 2 SGGS--NIKVVVRPFPNAREIDGKACIVNME--GNQITLTPPGAEERKARKSGRTIM 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 AGGSESVKVVVRCPRLNGEKADGRSRIVDMVDQGVKVRPNKADASEP----- 54

QY 57 DGPKAFAPDRSY-WSFDKNAPNARYAQEDLFDQLGVPLLDNAFGKYNCAIFAYGTGSGKS 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 --PKAFTFDQVYDWN-----CQQRDVFDTARPLIDSCIEGNGTIFAYGTGTGKS 104
```

```
QY 116 YSMMGYKE---HGVIPRICODMFRINELQKKNLTCTVEVSYLEIYNERVRLDLPST 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 HTMGKDEPPRLGLIPNTFYFVEITARDSGTKEF--LVRSYLEIYNEEVRLLGKDH 162

QY 173 KGNLKVREHPSTGTPYEDLAKLVVRSFQEIENLMDGKNKARTVAATNNNETSSRHAVFT 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 SKKMWELKESDPGVYVVDLSQFVCKNYEEMKNKVLGADNRQVGTATLMNQDSSRSHSIFT 222

QY 233 LTL-----TQKHDEETKMDTEKVAKISLVDLAGSRATSTGATGARLKEGABIN 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 ITIEICIEKLESAQAQPKAKKDDSNHVRVGLNVLVDLAGSERQDKTGATGDRLEKIGIN 282

QY 283 RSLSTGLRVTAALADMSSGQKQKQLVPYRDSVLTWLLKDSLGGNSMTAATAISPADIN 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 LSLTALGNVTSALVDGKSGH-----IPYRDSKUTRLDQLDGLGGNTKTVMVANIGPADWN 336

QY 343 FEETLSTLRVADSARKTKNHAVVNEPDNARMIRELKEELAAQLRSKLQSSGGGGG---A 398
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 YDEPMSLYRANRAKNTQNKPKINEDPKDMLRQFQEEIKKLEQLAARAAGGGGPTMP 396

QY 399 GSGGPGVE-----ESYPPD-----TPLEKOIVSIQDPATVYKMKSAEIVEQLNQ 443
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 SGGSPTQKIVERTEEVDPDIDAIAQMAEAKMKRSDISTEALDKAREEAATAAK-KQ 455

QY 444 SEKLYRDLNQTWEEKLAKTEEIHKERERAALEELGISIEK 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 LQAIIDQDGKTEAQKKAARDALKQAEEARAAG-AIEK 493

RESULT 14
A47334
Lckin kinesin-related antigen - Leishmania chagasi (fragment)
C:Species: Leishmania chagasi
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: A47334
R:Burns Jr., J.M.; Shreffler, W.G.; Benson, D.R.; Ghalib, H.W.; Badaro, R.; Reed, S.G.
Proc. Natl. Acad. Sci. U.S.A. 90, 775-779, 1993
A:Title: Molecular characterization of a kinesin-related antigen of Leishmania chagasi
A:Reference number: A47334; MUID:93133867
A:Accession: A47334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-955 <BR>
A:Cross-references: GH:L07879; NID:9308884; PIDN:AAA29254.1; PID:9308885
A:Experimental source: MHOM/BR/82/BA-2,C1
A:Note: sequence extracted from NCBI backbone (NCBIN:122864, NCBIPI:122865)
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; p-loop
F:13-398/Domain: kinesin motor domain homology <KMOT>
F:122-129/Region: nucleotide-binding motif A (P-loop)

Query Match 20.2%; Score 814; DB 2; Length 955;
Best Local Similarity 28.1%; Pred. No. 1.8e-36;
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

QY 7 IKVVVVRPFPNARE--IDRGAKCIVRMENQNTILTPP-----PGAEEKARKSGKTIMD 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 VKSVVRVRLPLNERENNAPEGTKVTVAAQAAAATVTVKVLGGSNNNGAAESMGTRARVAQD 72

QY 58 GPKAFADRSYWSF---DKNAPNARYAQEDLFDQLGVPLLDNAFGKYNCAIFAYGTGSGSK 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 ----QFDFHVFWSVETPDACGATPATQADVFTTGYPLVQHAFDGFNSCLFAYGQTGSGK 128

QY 115 SYSMMG-----YKHEGVIPRICODMFRINELQKKNLTCTVEVSYLEIYNERVRLD 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 TYTMGADVLSALSEGNGVTPRICLEIFARKASVEAQAQHSRWIVELGYVVEYNERVSDLL 188

QY 169 NPSTKG-----NLKVREHPSTGTPYVEDLAKLVVRSFQEIENLMDGKNKARTVAATNM 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 GKRRKGVKGGGEEVYVDVREHPSRGVFLQGRQLVEVGLDQVRLIEITGNGVRHTASTKM 248
```

```
QY 221 NETSSSHAVFTLTL-----TQKHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLK 276
Db 249 NDRSSRSIAIIMLLREERTMTTKSETIRTPACKSSRMNLVLDLAGSERVAQSQVEGQQPK 308
QY 277 EGAEINRSLSTLGRVIAALADSS--GKOKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAA 335
Db 309 EATHINLSITLGRVIDLADMATKAKAQSVPAPRDSKLTFFILKDSLGGNSKTTWIAAT 368
QY 336 ISPADINFEETLTLRYADSAKRIKHAVNEDPNARMRELKELAQLSKLSQSGGG 395
Db 369 VPSALNUEETLTLRYASARDIVNVAQVNEPRARRILELEQMEDMROAM-----421
QY 396 GGAGGGGPVEE-----SYPPDTPLEK 417
Db 422 --AGGPAYVELKKLALLESAQKRAADLOALEREREHNOVQERLLRATEAKSELES 479
QY 418 QIVSTQO-----PDATVKKMAEIVEQL 441
Db 480 RAAALQEMTATRRQADKQOALNLRKEEQARKERELLKEMAKDAALSKVRRKDAEIA 539
QY 442 NOSEKLYRLDQWEEKLAKTEEHKEREAAELGISIEKFGVGPYHSEMPHLYNLS 501
Db 540 SERELK-----ESTVAQLEREQREVALDAL-----OTHOKLQEALESSE 581
QY 502 -----DPLAECLVYNIKPGOTRVGNVQNTQAEIRLNGSKILKEHCTFENVDNVTIV 555
Db 582 RTAERKDLQOQ-----LTELSEKRTQLSQVVTDRERLTR-DLQRIQYEGETELARDVALCA 638
QY 556 PNB-----KAAMVNGVRIKDKPTLRSLGRYRIILGDFHIFRPNHPEEARERQBSLLRHS 610
Db 639 AQEMEARYHAAVFLQTLLELATEMEDALR-----ERALAERDEAAAAE--682
QY 611 VTNSQLGSPAGRDRTLSKAGSDAGDSRSDSPLPHFRGKSDWFFYARREASATILGID 670
Db 683 -----LDAALSTON-----ARESACERTLSLE 705
QY 671 QKISHITDDELALFDVOKARVRRGIVEDNEDSDSQSPFVRDKYMSNGTIDNFSLT 730
Db 706 QOLRE--SPERAELASOLEATAAAKSSAQDRENTATLEQOLRES-----EARAEL 757
QY 731 AITMPTGTRSDGDDGALFFGDKKSKODASNV-----DVEELRQQAQMEALKTAK 781
Db 758 ASOLEATAAA-----KMSAQDRENTATLEQOLRDSEERAAELASOLESTTAAK 807

RESULT 15
B44259
kinesin-related protein KIF3A - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 16-Jun-2000
C:Accession: B44259; S27872
R:Alzawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.
J. Cell Biol. 119, 1287-1296, 1992
A:Title: Kinesin family in murine central nervous system.
A:Reference number: A44259; MUID:93077686
A:Accession: B44259
A:Molecule type: mRNA
A:Residues: 1-701 <ATZ>
A:Cross-references: EMBL:D12645; NID:9220469; PIDN:BAA02166.1; PID:g220470
A:Experimental source: Brain
A>Note: sequence extracted from NCBI backbone (NCBIP:118911)
C:Complex: heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer associates with
C:Function:
A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal tra
C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; P-loop
F:1-368/Domain: head globular #status predicted <HGL>
F:15-351/Domain: kinesin motor domain homology <KNOT>
F:100-107/Region: nucleotide-binding motif A (P-loop)
F:369-599/Domain: helical rod #status predicted <ROD>
F:600-701/Domain: tail globular #status predicted <TGL>
F:106/Binding site: ATP (Lys) #status predicted
```

```
Query Match 19.9%; Score 800; DB 1; Length 701;
Best Local Similarity 31.8%; Pred. No. 6.5e-36;
Matches 231; Conservative 110; Mismatches 231; Indels 154; Gaps 25;

QY 6 NIKVYVVRPFPNAREID--RGAKCIVRMEGNOTILTPPGAEEKARKSKTKTMDGPKAF 62
Db 14 NVKVVVRCRLNEREKSCYQAVSVDEMRTITV-----HKTDSSN-----EPKTF 61
QY 63 AFDRSYNSFDKNAPYARQEDLFODLGVPLDNAPFKYNNCIFYAGQTSKSKSYMMGYG 122
Db 62 FDTIVF-----GPE--SKOLDVYNLTARPIIDSVLEGYNGTIFAYGQTGTGKTFMEGVR 114
QY 123 KE--HGVIPIRCQDMFRINELOKNLCTCTVEVSYLEIYNERVROLLNPSTKGNLKV 179
Db 115 AVPGIRGVIPNSFAHIFGHIAKAGDTRP--LVRSYLEIYNEVRDLIGKDOTORLEVK 172
QY 180 EHPSTGYVEDLAKLVRSFOBIENLMDENKARTVAATNNNETSSSHAVFTLTL--TO 237
Db 173 ERPDVGVIYKDLISAYVYVANNADMDRIMTLGHKNRSVGATNNNEHSSRSHAIPTIECSE 232
QY 238 KWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAAL 297
Db 233 KGVGNHVV--RMGKLHLVLDLAGSERQAKTGATGQRLKEATKINLSLTGNVISALVD 289
QY 298 MSSGOKKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLTLRYADSAK 357
Db 290 -----GKSTHVPYRNSKLTRLQDSLGGNSKTMCMANIGPADYNYDETISTLRYANRAK 343
QY 358 RIKNHAVNEDPNARMIRELKEELAQLRSKLOSSGG--GGGAGGGSGGPVEE-----407
Db 344 NIKKARINEDPKDALLQFOKBEIEELKKLEGEVSGSDISGSEDEDEGELEDGEK 403
QY 408 -----SYPPDTPLEKQ-----IVS 421
Db 404 RKRRDQAGKKYSPDKMVENQAKIDERKALETKLDMEEERKARAELERRERKDLKA 463
QY 422 IQQPDATVYKKMSKAE-----IVEQLNQSEKLYRLDNLQWEEKLAKTEEHKEREAA 472
Db 464 QEHQSLEKLSALEKKYIVGVVDLLAKAEQEKILLEESNMEELEERRRRAEQRLKELEEK 523
QY 473 -----LEELGISIEKFGVGPYHSEMPHLYNLSDDPLLAELCLVYNIKPGOTRVGNVQD 526
Db 524 EOERLDIEEKYTSLOEEAOG--KTKKLKVVWTM-----LMAAKSEMADLQOE 568
QY 527 TQAEIR--LNGSKILKEHCTPEN--VDNVV-----TIVPN-----EKAAMV 564
Db 569 HQRETEGLENIRQLSRELRLQMLIIDNFTPDYQEMENYVHWNEDIGEWOLKCVATG 628
QY 565 NGVRIDKPT---RLBSGYRIILGDFHIFRPNHPEEARERQBSLLRHSVTNLSQLGSPAP 621
Db 629 NNMKQTPVPDKRKDPFEVDLS--HVV-LAYTEESL--ROSLMKLERPRYSKGRPKPT 683
QY 622 GRHDT 627
Db 684 GRKRS 689
```

Search completed: April 25, 2001, 10:14:07
Job time: 283 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:13:34 ; Search time 26.31 seconds
(without alignments)
1020.764 Million cell updates/sec

Title: US-09-235-416-1

Perfect score: 4030

Sequence: 1 MSGGNIKVVVRVFPNARE.....ELRQQQAQMEALKAKQEF 784

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1668.5	41.4	1816	1	KF1B.HUMAN	O60333 homo sapien
2	1667	41.4	1103	1	KF1C.HUMAN	O43896 homo sapien
3	1663.5	41.3	1690	1	KF1A.HUMAN	Q12756 homo sapien
4	1663.5	41.3	1695	1	KF1A.MOUSE	P33173 mus musculus
5	1660.5	41.2	1816	1	KF1B.MOUSE	O60575 mus musculus
6	1617	40.1	1097	1	KF1D.RAT	O35787 rattus norv
7	1593.5	39.5	1584	1	U104.CABEL	P23678 caenorhabdi
8	1554.5	38.6	689	1	KF1B.RAT	O88658 rattus norv
9	860	21.3	699	1	K122_STRPU	P46872 strongyloce
10	848.5	21.1	747	1	KF3B.HUMAN	O15066 homo sapien
11	844.5	21.0	747	1	KF3B.MOUSE	O61771 mus musculus
12	827.5	20.5	742	1	K121_STRPU	P46871 strongyloce
13	825.5	20.5	786	1	F110_CHLRE	P46869 chlamydomon
14	814	20.2	955	1	K1NL_LEICH	P46865 leishmania
15	800	19.9	701	1	KF3A.MOUSE	P28741 mus musculus
16	789.5	19.6	702	1	KF3A.HUMAN	O94496 homo sapien
17	782.5	19.4	1231	1	KF4A.MOUSE	P33174 mus musculus
18	772	19.2	928	1	KF1H_NEUCR	P48467 neurospora
19	770.5	19.1	796	1	KF3C.MOUSE	O35066 mus musculus
20	769.5	19.1	796	1	KF3C.RAT	O55165 rattus norv
21	765.5	19.0	793	1	KF3C.HUMAN	O14782 homo sapien
22	758	18.8	1232	1	KF4A.HUMAN	O95239 homo sapien
23	746.5	18.5	784	1	K168_DROME	P46867 drosophila
24	743	18.4	935	1	K1NH_SYNRA	O43093 synccephalas
25	708	17.6	672	1	OSM3_CABEL	P46873 caenorhabdi
26	691.5	17.2	963	1	K1NH.HUMAN	P33176 homo sapien
27	691	17.1	963	1	K1NH_MOUSE	O61768 mus musculus
28	687.5	17.1	1031	1	K1NH_STRPU	P35978 strongyloce
29	684.5	17.0	1032	1	K1NH.HUMAN	Q12840 homo sapien
30	684	17.0	967	1	K1NH_LOLPE	P21613 loligo peal
31	676.5	16.8	1027	1	K1NN_MOUSE	P33175 mus musculus
32	673	16.7	1057	1	EG5_HUMAN	P52732 homo sapien
33	671.5	16.7	815	1	K1NH_CABEL	P34540 caenorhabdi

34	670.5	16.6	957	1	KF5C_HUMAN	O60282 homo sapien
35	669	16.6	956	1	KF5C_MOUSE	P28738 mus musculus
36	668	16.6	2663	1	CENE_HUMAN	Q02224 homo sapien
37	649.5	16.1	975	1	K1NH_DROME	P17210 drosophila
38	634	15.7	1056	1	K125_ARATH	P82266 arabidopsis
39	633	15.7	1184	1	B1MC_EMENI	P17120 emericella
40	631.5	15.7	883	1	YB3D_SCHPO	O14343 schizosacch
41	630	15.6	1067	1	EG52_XENLA	Q91783 xenopus lae
42	626.5	15.5	1111	1	K1P1_YEAST	P28742 saccharomyc
43	623	15.5	1060	1	EG51_XENLA	P88025 xenopus lae
44	621.5	15.4	1006	1	K125_TOBAC	O33826 nicotiana t
45	610.5	15.1	1066	1	K161_DROME	P46863 drosophila

ALIGNMENTS

```

RESULT 1
KF1B_HUMAN
ID AC O60333;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE KINESIN-LIKE PROTEIN KIF1B (KLP).
GN KIF1B OR KIAA0591.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bougueret L., Dufaur-Gare I., Grel P.;
RT "DNA encoding a kinesin-like protein (hklp) comprising biallelic
RT markers.";
RL Patent number WO0063375, 26-OCT-2000.
RN [2]
RP SEQUENCE OF 479-1816 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE OF 1449-1816 FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Salto K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SURFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AX039604; CAC16629.1; -

```



```
KW Motor protein; Microtubules; ATP-binding; Coiled coil;
FT Phosphorylation.
KW DOMAIN 1 355 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 359 388 COILED COIL (POTENTIAL).
FT DOMAIN 438 479 COILED COIL (POTENTIAL).
FT DOMAIN 523 590 FHA.
FT DOMAIN 633 674 COILED COIL (POTENTIAL).
FT DOMAIN 828 872 COILED COIL (POTENTIAL).
FT NP_BIND 97 104 ATP (POTENTIAL).
FT CONFLICT 669 689 Q -> R (IN REF. 2).
FT CONFLICT 956 962 LMGGRGV -> SGGGGGL (IN REF. 2).
FT CONFLICT 976 977 NV -> KL (IN REF. 2).
SQ SEQUENCE 1103 AA; 123071 MW; F14BC398D4B2EC10 CRC64;

Query Match 41.4%; Score 1667; DB 1; Length 1103;
Best Local Similarity 47.5%; Pred. No. 9.6e-80;
Matches 354; Conservative 121; Mismatches 157; Indels 114; Gaps 16;

QY 4 GGNKVVVRPNAREIDRGAKCIVRMENQNTILTPPPGAEEKARKSGKTMIDGPKAFA 63
DB 3 GASVAVVRPNARETSQDAKCVVSMOQNTTSIINP-----KQSKDAPKSET 51
QY 64 FDRSYWFDKNA--NYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTGSGKSYMWVG 122
DB 52 FDYSYWSHTSTEDPQASQOQVVRDIEGEMLLHAFEGYVNCIFAYGQTGAGKSYTHMGRQ 111
QY 123 K--ENGVIPTICQDMFRINELOKDKNITCTVYSLEYINERVDLLNPSTKGNLKVRE 180
DB 112 EPQOQGVLPQCELFDSRVSENG--SAQLSYSVESYMEIYCERVDLLNPKSRGSLRVE 170
QY 181 HPSTGPYVEDLAKLVKRSQETENLMDGNGKARTVAATNMNSTRSHAVFTLTTLQKWH 240
DB 171 HPILGPYVDLSKLAVTSYADIALMDGNGKARTVAATNMNSTRSHAVFTLVTFQRCH 230
QY 241 DEETKMDTEKVAKLSIVDLASERATSGATCARLKEGAEINRSLSITLGRVIAALADMS 300
DB 231 DQUTGLDSEKYSKISIVDLASERADSSGAGMRLEKAGANINKSLTLGKVISALADWQS 290
QY 301 GKOKNQLVPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSLRYADSAKRK 360
DB 291 -KKRSDFIPYRDSVLTWLLKENLGNSTRTAMIAALSPADINYEETLSLRYADRTKQIR 349
QY 361 NHAVYNEDPNARMTRELKELAQLSKLSQSGGGGGAGG----- 400
DB 350 CNAIINEDPNARLIRELQGEVARLRELLMAOGLSASALEGLTEGSGRGALPAVSSPPA 409
QY 401 -----SGPVEESYPPDTPLEKQIVSTQPDATVKKMSKAEIYEQLNQSEKLYRDL 451
DB 410 PVSPTSPTTHNGELSPFSPT--ESQI-----GPEAMERLQETEKIAEL 454
QY 452 NOTWEKLAETETHEKREALEELGISTEK--GFVGPVYHSEKPHLVNLSDDPLLAEL 509
DB 455 NETWEKLRKTEARWERALEAMGVAVRDGTVGVFSPAKTPHLVNLSDDPLMSECL 514
QY 510 VYNIKPGQTRVGNVQDQAEIRLNGSKLKEHCTFENV-----DNVVTIVPNEKAAMV 564
DB 515 LYHIKDGTVTRVGQVDMQ----IKLTGQPIREQHCLFRSIPQPDGEVVVTLPECEGAETV 570
QY 565 NVGRIDKPTLRSGYRITLIGDFHFRFNHPEARARQRLSLRHSVTNSQLSGPAGRH 624
DB 571 NGKLVTEPLVLKSNRIVMGKNVFRFNHPEARLER-----GVPPP--- 614
QY 625 DRTLSKAGSDAGDSRDSPLPHFRGKDSDMFYARREASATILGDUKQISHLTDDELDA 684
DB 615 -----PGPPEPDWNEFAQKELLEQ--QGTDIKLE--MEKRLQDL 650
QY 685 FDDVOKARVRGLVEDNE-----DSDS 707
DB 651 ENQYRKEKEEADLLLEQQQLYADSDS 676

RESULT 3
```



```
Db 229 RHDAETNITTEKVKISLVLAGSERADSTGAKGTRLKEGANINKSLTLGKVISALAE 288
QY 299 SSG-----KQKNQVLPPYRSVITWLLKDSGGNSMTAMTAISPADINFEETLSLRYA 353
Db 289 DSGPNKKKKKKTDPYRDSVITWLLKDSGGNSRTAMVAALSADINFEETLSLRYA 348
QY 354 DSARKIKHAVNEDPNARIRELKEAQLSKLSSGGGGGGGGGGGGGGGGGGGGGGGGGG 413
Db 349 DRAKQIRCNALINEDPNKKILRELKDEVRDLRYAAGLG-----DIT 392
QY 414 PLEKQIVSLQPDATVKKMS-----KAEIVQOLNSEKLYRDLNQW 455
Db 393 DMTNALVGM-SPSSSLGALSRAASVSLHERILFAPGSEAEIKETIKETIAELNETW 451
QY 456 EEKLAETIEHKEREAALEELGISTEK--GFGVGYHSHKEMPHLVNLSDDPLLAARCLVYNI 513
Db 452 EEKLRRTETAIREREALLMGVAMRGDGGTLGVFSPKTPHLVNLNEDPLMSSECLLYI 511
QY 514 KPGOTRVGNVODTQAEIRLNGSKILKEHCTFEN-----VDNVVIVPNEKAAVWNGVR 568
Db 512 KGVTVRGREDERRQDVLGSHFKEEHCIFRSRDSRGSGGEAVVLEPCGADTYVNGKK 571
QY 569 IDKPTRLRSYRIILGDFHIFRNHPPEARAEQESLLRHSVTSQLSGSPAGRGRDRTL 628
Db 572 VTEPSILRSGNRIIMGKSHVFRNHPQARQER----- 606
QY 629 SKAGSDADGSRSDPLPHFRGKDSWYARREASAILGLDQKISHITDDELDALEPDY 688
Db 607 -----TPCAETPAPVDMAFQRELLEK-QCIDMK--QEMEQRIQLEDDY 649
QY 689 OKARAVRGLVEDNEDSDSSPPVDKYMEN 720
Db 650 RREREATVLE-QORLDYESKLEALQKQMS 680

RESULT 5
KF1B_MOUSE STANDARD; PRT: 1816 AA.
AC Q60575; Q9WVE5; Q9R0B4; Q92119;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Brain;
RX MEDLINE=95094296; PubMed=7528108;
RA Nangaku M., Sato-Yoshitake R., Okada Y., Noda Y., Takemura R.,
RA Yamazaki H., Hirokawa N.;
RT "Kif1b, a novel microtubule plus end-directed monomeric motor protein
RT for transport of mitochondria.*";
RL Cell 79:1209-1220(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=ICR; TISSUE=Brain;
RX PubMed=10571041;
RA Gong T.L., Winnicki R.S., Kohrman D.C., Lomax M.I.;
RT "A novel kinesin of the UNC-104/Kif1 subfamily encoded by the Kif1b
RT gene.*";
RL Mamm. Genome 10:0-0(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,
RA Perry V.H., Coleman M.P.;
RT "The major brain isoform of Kif1b lacks the putative mitochondria-
RT binding domain.*";
RL Mamm. Genome 10:0-0(1999).

[4]
SEQUENCE FROM N.A. (ISOFORM 2).
RP STRAIN=ICR;
RA Nakagawa T., Hirokawa N.;
RT "Identification and characterization of a new kinesin superfamily
RT Kif1b-beta.*";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC MICROTUBULE PLUS END-DIRECTED MOTILITY.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESTICLE- AND
CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/BETA AND 3;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
CC TISSUE (MAINLY IN THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE
CC TYPE OF NEURONAL CELL.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D17577; BAA04503.1; -.
EMBL: AF090190; AAF06718.1; -.
EMBL: AF131865; AAD39438.1; -.
EMBL: AB023656; BAA75243.1; -.
HSSP: P33176; 1BG2.
MGD: MGI:108426; Kif1b.
DR InterPro; IPR000253; -.
DR InterPro; IPR001752; -.
DR InterPro; IPR001849; -.
DR Pfam; PF00438; FHA; 1.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil;
KW Alternative splicing.
FT DOMAIN 1 361 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 365 386 COILED COIL (POTENTIAL).
FT DOMAIN 470 502 COILED COIL (POTENTIAL).
FT DOMAIN 556 612 FHA.
FT DOMAIN 668 737 COILED COIL (POTENTIAL).
FT DOMAIN 841 869 COILED COIL (POTENTIAL).
FT DOMAIN 1702 1799 PH.
FT NP_BIND 97 104 ATP (POTENTIAL).
FT VARSPIC 289 294 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPIC 394 434 IDPLIDDYSGSGYKLVKFNKKHYLLASENORPQNFSTA
FT 707 1196 -> T (IN ISOFORM 2 AND ISOFORM 3).
FT YESQLAORQVETRSIAAETEEEEEEEEEEVPTWTFEFELA
FT QWAFRWKSHQFTSLRLDGMNAVYLLKANAISVELKKVQ
FT QFVLLTDTLSPVPPPELLPSEMEXTHEDEPPRTVAVVE
FT QDLKNGATHYWSLKLKRLDLRMRYDRAGEVASSAQDPS
FT ETTMTGSDPFYDRPHMFKLVGSSPIPHGCVNERLADRTSP
FT TFSADSDITELADEQDAMEDFDDEAFVDVDTGSDAGTEG
FT SELFSDGHDPEDVSRPMFLVGLVGRVLSNLLVPVLIHRV
FT AIYSEKGEVGRFLVAVQATADEAPDYSGIRQSGTARI
FT PDENEYFNQSDFSAAATRSGLSLSEELRIEVEGQSQSEVIS
FT PPEVNRNMNDLKLKSLDGLDKMWMEGFSEIGNHLKLGS
FT FTRFVTLQASGILPEYADIFCQFNFLHRHDEAFSTELKN
FT NGRGSPFAGYHVQNIATEVTEFVDYIKTRPIVFEVFGH
FT -> ADSDSGSDSKRSCESKWLITSLREKLPPSKLQITVQ
FT KGLPSSGKKREPIKMYQIPQRRRLSKDSKWTISLKLQVQ
```

```
FT VKETCYEVALNDRFHRHQEIEALAIKMKELCAMYKKDPN
FT ERDSRAVARDVMDTGVGDEKEDMMVTGKGGTDDVLLKV
FT HIDEKLEILOEVKQNNMKDDEIKVLKMKELKVLPLIG
FT SQEKSOGSHKTEPLVAGANSYSDNGYSKGESELGKEER
FT VSQKMGDPAPFRKRLWRKROEQIRFKNLQOQELTKOLRRO
FT NVPHRFPPNRPFRPFKPNKPHNSWSPGTHIITEDEV
FT IELRIPKDEKESQEKYGRRAASRDVQSWAGTTRSD
FT HIOVSKOHSNQOPPOLRWRKNSLNNGQPKTRCOATASS
FT ELSNHSHPADLOTFOAKRHJHOHRQPCVNTYNGQVRS
FT STASCCOKDQKSHCNQVTPPPMRKQFSAPNLKAGRETT
FT V (IN ISOFORM 3).
FT MISSING (IN ISOFORM 3).
FT G -> V (IN REF. 1 AND 2).
FT GGTG -> RGD (IN REF. 1).
FT P -> S (IN REF. 4).
FT KL -> TW (IN REF. 3).
FT I -> V (IN REF. 3).
FT D -> G (IN REF. 4).
FT SQ SEQUENCE 1816 AA; 204080 MW; E316EC295138E5DE CRC64;

Query Match 41.2%; Score 1660.5; DB 1; Length 1816;
Best Local Similarity 45.1%; Pred. No. 4.1e-79;
Matches 357; Conservative 127; Mismatches 170; Indels 137; Gaps 17;

Qy 4 GGNITKVVVRVPPENAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTMIDGPKAFA 63
Db 3 GASVKVAVRVPFNSREFTSKESKCIQMGNSTSIINPKPK-----APKSFS 51

Qy 64 FDRSYWSP-DRNAPNYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTGSGKSYMMGYG 122
Db 52 FDSYWSHTSPEDCFASQNRVNDIGKEMLLHAFEGYVNCIFAYGQTGAGKSYTMGKQ 111

Qy 123 KEH--GVTPRICQDMFRINELOQDKNUTCTVEVSYLYIYNERVLDLNPSTKGNLKVRE 180
Db 112 BESQAGIIPOLCEELFEIND-NCNEEYSYSEVSYMEIYCERYVDRDLNPKNGKLNRYRE 170

Qy 181 HPSTGPYVEDLAKLVRSFOETENLMDGKNARTVAATNMNETSSRSRAVFTLTITOKWH 240
Db 171 HPLGYPVEDLSKAVSYTDIADMDAGNARTVAATNMNETSSRSRAVFTIYFTOKKQ 230

Qy 241 DEETKMDTEKAKISLVDLAGSRATSTGATGARLKEGAEINRSISTLGRVIAALAD-- 298
Db 231 DPETNLSTKYSKISLVDLAGSRADSTGATGTRLKEGANINKSLTTLGKVISALAEVDN 290

Qy 299 ---SSGKOKKQOLPYRDSVLTLLKDSLGNSWTAMTAALSPADINPEEPLSTLRYADS 355
Db 291 CTSKSKKKKKTDFIPYRDSVLTLLRENGLGNSRTAMVAALSPADINDETSLTRYADR 350

Qy 356 AKRIKNAHVVEDPNARMIRELKEELAOLRSKLOSSGGG-----GGAGG--- 400
Db 351 AKQIKNAVINEDPNAKLVRELKEEVTRLKOLLRAQGLGDIIDIDPLIDYSGSGGKYLK 410

Qy 401 -----SGGVEESYPPDTPLEKQ-----IVSIQQPDATYKK 431
Db 411 DFQNNKHYLLASENQRPNGFNSTASMGSLTSS-PSSCSLNSQVGLTSVTSIQ--ERINST 467

Qy 432 MSKAIIVQLNOSEKLYKDLNOTWEEKLAKTEETHKEREAALEELGISIEK--GFVGYPH 489
Db 468 PGGEAIERLAKSEKIIAELNETWEEKLRKTEATNMRERALLAEMGVAIREDDGGTGLVFS 527

Qy 490 SKEMPHNLNDDPLLAELCLVYNIKPGOTRYGVNODTQAEIRLNGSKILKEHCTFENV- 548
Db 538 PKKTHLNLNEDPLMSCLLYIKDGLTRVGQADARRQDVLGSAHIKEEHLFRSER 587

Qy 549 ----DNVYTIYPNKAAYVWNGVRIDKPTRLSRGVRIILGDFHIFRPNHPEEAAREQEO 604
Db 588 SNTGEVIITLPECSESYVNGKRAVHPQLRSGNRIIMGNKHVFRFNHPQARAREK- 646

Qy 605 SLLRHSVTNSQLGSPAGRHDRDTLSKAGSDADGDSRSDPLPHFRGKDSDFYARREAS 664
Db 647 -----TPSAETPSEPVDWTFQAQRELLE 668
```

```
Qy 665 AILGLDOK-----ISHLTDDELDA-----FDDVQKARAVRRGL-----VED 701
Db 669 K-QGIDMKQEMEKRLQEMEILYKKEEADLLQLQORLDYESKLQALQROVETRSAAET 727

Qy 702 NEDSDSSQSF 712
Db 728 TEEEEEVEVP 738

RESULT 6
KFID_RAT ID KFID_RAT STANDARD; PRT; 1097 AA.
AC O35787;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIFID.
GN KIFID.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RA Rogers K.R., Griffin M., Brophy P.J.;
RT "The secretory epithelial cells of the choroid plexus employ a novel
RL kinesis-related protein."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ000696; CAA04248.1; -.
CC HSP; P17119; 3KAR.
CC InterPro; IPR000253; -.
CC InterPro; IPR001752; -.
CC Pfam; PF00225; kinesin; 1.
CC Pfam; PF00498; FHA; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
CC PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
CC PROSITE; PS50006; FHA_DOMAIN; FALSE_NEG.
CC Motor protein; Microtubules; ATP-binding; Coiled coil.
CC DOMAIN 1 357
CC DOMAIN 358 380 MECHANOCHEMICAL (MOTOR).
CC DOMAIN 437 478 COILED COIL (POTENTIAL).
CC DOMAIN 520 587 COILED COIL (POTENTIAL).
CC DOMAIN 630 671 COILED COIL (POTENTIAL).
CC DOMAIN 824 868 COILED COIL (POTENTIAL).
CC NP_BIND 96 103 ATP (POTENTIAL).
CC SEQUENCE 1097 AA; 122333 MW; 8EF40B1C7579BA5B CRC64;

Query Match 40.1%; Score 1617; DB 1; Length 1097;
Best Local Similarity 45.9%; Pred. No. 3.9e-77;
Matches 357; Conservative 129; Mismatches 184; Indels 108; Gaps 20;

Qy 4 GGNITKVVVRVPPENAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTMIDGPKAFA 63
Db 3 GASVKVAVRVPFNSREFTSQDAKCVVSMOQNTTSIIN-----KOSRFL---KA-S 50

Qy 64 FDRSYWSPDK-NAPNYARQEDLFQDLGVPLLDNAFKYNNCIFAYGQTGSGKSYMMGYG 122
Db 51 FDSYWSHTSPEDPQFASQOQVYRDIGEMLLHAFEGYVNCIFAYGQTGAGKSYTMGKQ 110
```

```
Qy 123 K--EHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRDLNPNSTKGNLKVRE 180
Db 111 EPQOQGIYVQICEDLFSRYN-VNQAQISYSEVSMEIYECERVDLNLNPKSRGSLURVE 169
Qy 181 HFTSTGPIYVEDLAKLVRSFOETENLMDGKNGKARTVAATNNNTSSRSHAVFTLTQKWH 240
Db 170 HPTLGPYVQDLSKLVSTYADIALDMCGNKARTVAATNNNTSSRSHAVFTVTFQRSH 229
Qy 241 DEETWMDTEKVAKISLVDLGASERATSTCATGARKLEGAENRSLSTLGRVIAALADWS 300
Db 230 DQTLGDSKVKISLVNLAGSERADSGARGMRLEKAGINKSLTTLGVKVISALADQS 289
Qy 301 GOKKNQIYVRSVLTWLLKDSLGNSMTAMIAAISPADINFEETLSTLRYADSARKIK 360
Db 290 -KKRSDFIPYRDSVLTWLLKNSGNSRTAMIAAISPADINFEETLSTLRYADRTKQIR 348
Qy 361 NHAVNEDPNARMIRELKEALQSKLQSSGGGGAGG-----SGG----- 403
Db 349 CNAYINEDPNARLIREQEVARLELLMAQGLSASALGKLVESGPGVLPFASSPPA 408
Qy 404 PVESYPPDTPLEKQIVSIQOPDATVTKKMSKAETVQLNQSEKLYRDLNQTWEELAKTE 463
Db 409 PASPSPP--PHNGELPSFSPSAE-POIGPEEAMERLQETEKIIAELNETWEELKURTE 465
Qy 464 ETHKREAALELIGSTENGKGVPHSKEMPHLVNLSDDPLLAELCLVYVYKPGQTRGVN 523
Db 466 ALRMEREAALLAENGSPGWRVTGVSPKTPHVLNLEDPLMSECLLYHKIDGVTRGV 525
Qy 524 NQDTAEIRLNGSLIKLHCHTENV-----DNVVTVPNEKAAVNVGVNVRIDKPTLRSG 578
Db 526 DVD----IKLTGQPIREQHCLFRSIPQDGEVMTLEPECEGAETVYNGKLVTPLEPLKSG 581
Qy 579 YRIILGDFHIFRNFHPEARAEQROSLRHVSNTSOLGSPAPGRDHRTLSKAGSDADG 638
Db 582 NRVMGKNHVFNFHPEQARLER-----GVPPP----- 611
Qy 639 SRSDSPHPHFGKDSWVFARREASAILGLDQIKSHLTDDELDAFDVQKARAVRGL 698
Db 612 -----FGPSPVDVMAQDKLEQ-OGIDIKLE--MEKRLQDLENQYRKEKEADLL 661
Qy 699 VE-----DNDSDSS-----FVPRKYMNSGTIDNFSLOTATIMPTGTPRS 740
Db 662 LEQRIYADSDGSDSKRSCEENRSLSSRLDELPPN-----TVQTIYKRCGLPSS 713

RESULT 7
U104_CABEL STANDARD: PRT: 1584 AA.
AC P23678;
DT 01-JUN-1991 (Rel. 20, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN UNC-104.
GN UNC-104.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91097805; PubMed=1846075;
RA Otsuka A.J., Jeyaprakash A., Garcia-Anoveros J., Tang L.Z., Fisk G.,
RA Hartshorne T., Franco R., Born T.;
RT "The C. elegans unc-104 gene encodes a putative kinesin heavy
RT chain-like protein."
RL Neuron 6:113-122(1991).
RC -I- FUNCTION: INVOLVED IN MICROTUBULE-ASSOCIATED TRANSPORT.
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC -I- SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M58582; AA03517.1; .
DR PIR: JN0114; JN0114.
DR HSP: P56536; 2KIN.
DR InterPro: IPR001752; .
DR InterPro: IPR001849; .
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00225; kinesin; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Transport.
FT DOMAIN 1 354 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 425 445 COILED COIL (POTENTIAL).
FT DOMAIN 598 652 COILED COIL (POTENTIAL).
FT DOMAIN 777 797 COILED COIL (POTENTIAL).
FT DOMAIN 183 335 MICROTUBULE-BINDING.
FT DOMAIN 1460 1558 PH.
FT DOMAIN 957 1052 ARG/LYS-RICH (BASIC).
FT DOMAIN 1203 1584 ARG/LYS-RICH (BASIC).
FT NP_BIND 93 100 ATP (BY SIMILARITY).
FT VARIANT 598 598 I -> T.
FT VARIANT 930 930 V -> M.
SQ SEQUENCE 1584 AA; 179736 MW; 559462B3FD029B43 CRC64;
```

```
Query Match 39.5%; Score 1593.5; DB 1; Length 1584;
Best Local Similarity 44.7%; Pred. No. 1.1e-75;
Matches 356; Conservative 139; Mismatches 213; Indels 89; Gaps 19;

Qy 6 NIKVYVRFNFNAREIRGAKCIVRMENGTITLTPPGAEKARKSGKTMDGPKAFAD 65
Db 3 SVKYAVRFPNQREISNTSKVLQVNGNTTI-----NGHSINKENFSNFD 50

Qy 66 RSYSPDKNAPYARQEDLFODLGVPLDQNAFGVNNCFAYGOTGSGKSYMMCYGK-- 123
Db 51 HSYSPARNDPHFTQKVQVEELGVEMLEHAEGINVCIFAYGOTGSGSYTMGKANDP 110

Qy 124 -EHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDRDLNPNSTKGNLKVREHP 182
Db 111 DEMGIIPRLCNDLIFARDN--NNDKDVQYSEVSMEIYECERVDLNLNPNSGNLRVREHP 169

Qy 183 STGPIYVEDLAKLVRSFOETENLMDGKNGKARTVAATNNNTSSRSHAVFTLTQKWHDE 242
Db 170 LLGPIYVDLTKMAVCYSHDNLMDGKNGKARTVAATNNNTSSRSHAVFTIVLTKRHCA 229

Qy 243 ETKMDTEKVAKISLVDLGASERATSTCATGARKLEGAENRSLSTLGRVIAALADWSGK 302
Db 230 DSNLDTKEKSHKISLVDLGASERANSTGAEQRLKAGANKSLTTLGLVSKLAPESTKK 289

Qy 303 OKKO--LVPRYSVLTWLLKDSLGNSMTAMIAAISPADINFEETLSTLRYADSARKIKN 361
Db 290 KSKNKGVIYPRYSVLTWLLRNLGNSKTMALAAISPADINFEETLSTLRYADRAKQIVC 349

Qy 362 NHAVNEDPNARMIRELKEALQSKLQSSGGGGAGGSGGPGVPEESYPPDTPLEKQIVS 421
Db 350 QAVYNEDPNAKLIRELNEEVYIKRHLKDKG-----IDVTD 385

Qy 422 IQQPDATVTKKMSKA-----EIVEQLNCKSEKLYRDLNQTWEELAKTEITHKEEALEEL 476
Db 386 VQETPGKHKKGKPLPAHVHWEQLEKQSEKLMAEIKGTWQGLIHTTEIRKQREELRDM 445

Qy 477 GIS-IEKG-FVGPYHSKEMPHLVNLSDDPLLAELCLVYVYKPGQTRGVNVDQAEIRLN 534
Db 446 GLACAEDGTTLGVSFPKPLPHLVNLEDPLMSECLLYHKIDGVTRGVNVDQAEIRLN 505
```

```
QY 535 GSKILKEHCTFENVNVVTIYPNEKAAVMVNGVRIDKPTRLSGYRIILGDFHLEFRNHP 594
Db 506 GEATLEHCEFINEGNVTLTKMKNASCYINGKQVTPPTVLHTGSRVTLGHEHVRVNDP 565
QY 595 EEARAEQEQSLLRHSVTNSOLGSPAPGRHRDRTLSKAGSDAGDSRSDSPLPFHRGKDS 654
Db 566 QEARQSRHNLAAIAEQPIDWYA-----QQELLDKQIDLKADMEK-----KMLEMESQ 614
QY 655 WFYARREASAILGLDQKISHLTDELDALEDDVQKARAVRR-----GLVEDNED--SD 706
Db 615 Y---RREKVE---LEQKMYHOT-REYESMIENLKQVDLAQSYISGGGSIWEGERMULTS 666
QY 707 SQSSFPVRDKYMSNGTIDNFSLDTAIMP-----GTPSRDDDDGDALEFGDKKSKQDASND 762
Db 667 SLLEFPPEELKWTSDQ--KRVVLLKAAIKRWYHQTFSVRDDLMGNAIFV-----KEANAIS 718
QY 763 VEELRQQAQMEELAKT 779
Db 719 VE--LKKKVQFOFALLT 733

RESULT 8
ID KF1B_RAT STANDARD; PRT; 689 AA.
AC O88658;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B (FRAGMENT).
GN KIF1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN R1
RP SEQUENCE FROM N.A.
RA Faure K., Gruber D., Bullinski J.C.;
RT "Identification of kinesin-like molecules in myogenic cells.";
RL Eur. J. Cell Biol. 0:0-0(1998).
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AF083331; AAC33292.1; -.
DR HSSP; P33176; I8G2.
DR InterPro; IPR000253; -.
DR InterPro; IPR001356; -.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; kinesin; 1.
DR Pfam; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; FALSE_NEG.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 359 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 360 381 COILED COIL (MOTOR).
FT DOMAIN 451 483 COILED COIL (POTENTIAL).
FT DOMAIN 538 618 FHA.
FT DOMAIN 651 688 COILED COIL (POTENTIAL).
```

```
FT NP_BIND 97 104 ATP (POTENTIAL).
FT NON_TER 689 689
SQ SEQUENCE 689 AA; 77352 MW; F64818C12FB69319 CRC64;

Query Match 38.6%; Score 1554.5; DB 1; Length 689;
Best Local Similarity 47.0%; Pred. No. 3,7e-74;
Matches 335; Conservative 116; Mismatches 161; Indels 101; Gaps 17;

QY 4 GGNIKVVVRVPFNAREIDRGAKCIVRMENQOTILTPPPGAEAKARKSGKTIIMPGRKFA 63
Db 3 GASVKVAVRVPFNAREIDRGAKCIVRMENQOTILTPPPGAEAKARKSGKTIIMPGRKFA 63
QY 64 FDRSYWSEF-DKNAENYARQEDLFQDLGVPLDLNFAKGYNNCIFYAGQTGSGKSYSMGYG 122
Db 52 FDSYWSHTSPDEPCFASQSRVYNDIGKEMLLHAFEGYNNCIFYAGQTGAGKSYTMMGK 111
QY 123 KEH--GVIPRICODMFRINELQDKNLTCTVEYSYLIYNERVDRDLNLPSTKGNLKYRE 180
Db 112 EESQAGIIPTCCEELFEKIND-NCNEDMSYSVS-SYMEIYCERVDLLNPNKGNLRYRE 169
QY 181 HPSTGYPVEDLAKLVRSFQBIENLMDGCKARTVAATNMNETSSRSRAVFTLTITOKWH 240
Db 170 HPLLGPYVEDLSKLAVTSTYTDIADLMDAGNKARTVAATNMNETSSRSRAVFTIVFTQKK 229
QY 241 DEETKMDTEKVAKISLVDLAGSRATSTGATGARLKEGAEINRSLTGLRVIAALADMS 300
Db 230 DPETNLSTKVKTKISLVDLAGSRADSTGATGARLKEGAEINRSLTGLRVIAALADMS 289
QY 301 GKOKNOLVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETL--STLRVADSAR 358
Db 290 -KKKTDFIPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETL--STLRVADSAR 348
QY 359 IKNHAVNEDPNARMIRELKEELKSLQSSGGG----- 394
Db 349 IKCNVAVNEDPNARMIRELKEELKSLQSSGGG----- 394
QY 395 -GGAGSGGGPVEESYPDPDTLEKQ-----IVSQQDPATVKMKSKAEIVPOLNQSEKLY 448
Db 409 PGNFSTASMSGLTSS-PSSCSLNSQAGLTSTVSIQ--ERIMSTFGGGEAIBELKESEKII 465
QY 449 RDLNQTWEELKAKTEETHKEREALAEELGISIEK--GVGPGYHSEMPHLNLSDDPLLA 506
Db 466 AELNETWEELKAKTEETHKEREALAEELGISIEK--GVGPGYHSEMPHLNLSDDPLLA 506
QY 507 ECLVYNIKPGQTR-VGNVNODTQAEIRLNGSKILK-EHCTFENVDN-----VVTIVPNEK 559
Db 526 ECLLYYIKDGITKGFGQADAEERRQDIVLSGAHTIKEHCIFRSENNNTGEVITLPCER 585
QY 560 AAVMVNGVRIDKPTRLSGYRIILGDFHIFRNPHEPAREARQEQSLLRHSVTNSOLGSP 619
Db 586 SETVYNGKRVAAHPVOLKRSVYSSWVKNHVFRFNHPEQARERK----- 629
QY 620 AGRHRTLSKAGSDAGDSRSDSPLPFHRGKDSDFVAREEASAILGLDQK 672
Db 630 -----TPSAETPSEPVDWTFQAQRELLEK-QGIDMK 658

RESULT 9
ID K122_STRPU STANDARD; PRT; 699 AA.
AC P46872;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE KINESIN-II 85 KDA SUBUNIT (KRP-85/95 85 KDA SUBUNIT).
GN KRP85
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
```



```

QY 122 ---GKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEINERVRDLNPSKGNLKV 178
DQ 110 RGDPEKRGVINSDFHIFTHISRQNOQYL---VRASYLEYIQEIRDLNLSKDTQKRL 166
QY 179 REHSTGTYVEDLAKLVRSFQEIENLMDGKARTVAATNMNSETSRSHAVFTLTQK 238
DQ 167 KERPDGTGVVRLDSSFTVSKVEIHEVMVGNQNRSGVATNMNHESSRSHAFVITI--- 223
QY 239 WHDEETKMDTE---KVAKISLVDLGSRATSTGATCARLKEGAEINRSLTGLRVIAA 295
DQ 224 --ECSEVLGDCENHIRVGNKLVLDLAGSRQAQTGAOGERLKEATKINLSALGNVISA 281
QY 296 ADMSSGOKKQNLVYRDSVLTLLKSLGNSMTAMTAAISPADINFEETLSTLRVADS 355
DQ 283 VD-----GKSTHPIYRDSKUTRLQSLGNGNAKTVWVAVGPNVYVEETLTLRYANR 336
QY 356 AKRIKHAVVNEEDPNARMIRELKEALQRLSKLQSSGG-----GGAGSGSGPV 405
DQ 337 AKNIKNKPRVNEEDPKDALLREFQEEIARLKAQLEKRSIGRKRREKREGSGSGGGEE 396
QY 406 EESTPPTPLEKQIVSQPDATVKKSKAEIVEQ-----LNQSEKLYRDLNQTWEK 459
DQ 397 EEEGE-----EGEEGDDKDDYWRQEQEKLIEKRAIVEDHSLVAEEKM-----RLLEKE 447
QY 459 LAKTEIHKEREAALEELGISIE---KGFVGPY---HSEMPHLYNLSDDPLLAECV 511
DQ 447 EKMEDLRREKDAE-EMLGAKIKAMESKLLVGGKNIVDHTNEQOKILEQKQAEIAEQ--- 502
QY 511 YNIRPGQTRVGNVNDQTAQAEIRLNGSKILKEHCTFENVDNVVTIVPNEKAAMVNGVRID 570
DQ 503 -----KRREIEIQ--QMESRDEETLEKE--TYSSIQEQEVDI-----KTK 531
QY 571 KPTRLRSYRIILGDFHFRFNHPEBARAEQEQS-----LLRHSVTNSOLGSPAPGRH 624
DQ 540 KKLKLSKLAQVAAEIDHLEQEHKIQEQLQEQNELTLRLKHLITENFI--PLEEK- 596
QY 625 DRTLSKAGSDAGDSRSDSPHPHPRGKD-----SDWFYAR-----R 660
DQ 597 SKINWRAFFDEEDHUKLHPITRLNQQMKRPVSAVGKRPRLSQHARMGMIRPEARYR 656
QY 661 EAASAILGLQKISHLTDDELDFDDVQKARAVRRGLVEDNEDSDSQSS 711
DQ 657 AENIVLELMDPSRTTRDYEGPALAPRVQ--AALDAALQDEDELDQVDSF 705

RESULT 11
KF3B_MOUSE STANDARD; PRT; 747 AA.
AC Q61771;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN MOTOR 3B).
DE DE
GN KIF3B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=96032268; PubMed=7559760;
RA Yamazaki H., Nakata T., Okada Y., Hirokawa N.;
RT "Kif3A/B; a heterodimeric kinesin superfamily protein that works as a
RT microtubule plus end-directed motor for membrane organelle
RT transport.";
RL J. Cell Biol. 130:1387-1399(1995).
CC -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC ACTIVITY IN VITRO.
CC -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.

```

```

CC CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26077; BAA05070.1; -
DR HSP; P56536; 2KIN
DR MGD; MGI:107688; Kif3b.
DR InterPro; IPR001752; -
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT DOMAIN 1 345 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
FT DOMAIN 580 747 GLOBULAR.
FT NP_BIND 96 103 ATP (POTENTIAL).
FT DOMAIN 386 393 POLY-GLY.
FT DOMAIN 394 405 POLY-GLU.
FT DOMAIN 723 730 POLY-SER.
SQ SEQUENCE 747 AA; 85288 MW; FA369A4190EC8B47 CRC64;

```

```

Query Match 21.0%; Score 844.5; DB 1; Length 747;
Best Local Similarity 32.9%; Pred. No. 4.5e-37;
Matches 256; Conservative 116; Mismatches 253; Indels 153; Gaps 27;

QY 6 NIKVVVVRVRFENARE----IDRGAKCTVRMEGNOTILTPPGAEEKARKSGKTTMDGPKA 61
DQ 9 SVRVVVRCPWNGEKAASVDKVVVDVVKL--GVSVKNPKG-----TSEHPKPT 56
QY 62 FAFDRSY-WSFQKPNAPYARQEDLFDGLVPLLDNAPKYNNCIFAYGQTGSGKSYSNMG 120
DQ 57 FTDAVYDWN-----AKQFELYDETFRPLVDSVLQGFNGTIFAYGQTGKTGTYMEG 108
QY 121 Y---GKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEINERVRDLNPSKGNLKV 177
DQ 109 VRGPEKRGVINSDFHIFTHISRQNOQYL---VRASYLEYIQEIRDLNLSKDTQKRL 165
QY 178 VREHPSTGTYVEDLAKLVRSFQEIENLMDGKARTVAATNMNSETSRSHAVFTLTQ 237
DQ 166 LKERPDGTGVVRLDSSFTVSKVEIHEVMVGNQNRSGVATNMNHESSRSHAFVITI-- 223
QY 238 KWHDEETKMDTE---KVAKISLVDLGSRATSTGATCARLKEGAEINRSLTGLRVIAA 294
DQ 224 --ECSEVLGDCENHIRVGNKLVLDLAGSRQAQTGAOGERLKEATKINLSALGNVISA 281
QY 295 LADMSSGOKKQNLVYRDSVLTLLKSLGNSMTAMTAAISPADINFEETLSTLRVADS 354
DQ 282 LVD-----GKSTHPIYRDSKUTRLQSLGNGNAKTVWVAVGPNVYVEETLTLRYAN 335
QY 355 SAKRIKHAVVNEEDPNARMIRELKEALQRLSKLQSSGG-----GGAGSGSGPV 404
DQ 336 AKNIKNKPRVNEEDPKDALLREFQEEIARLKAQLEKRSIGRKRREKREGSGSGGGEE 395
QY 405 VE---ESYPPDTPLEKQIVSQPDATVKKSKAEIVE---LNQSEKLYRDLNQTWEK 459
DQ 396 EEEGEDEGEDDDKDDYWRQEQEKLIEKRA--IVEDHSLVAEEKM-----RLLEKE 447
QY 460 AKTEIHKEREAALEELGISIE---KGFVGPY---HSEMPHLYNLSDDPLLAECV 511
DQ 448 KKMEDLRREKDAE-EMLGAKIKAMESKLLVGGKNIVDHTNEQOKILEQKQAEIAEQ--- 502
QY 512 NIKPGQTRVGNVNDQTAQAEIRLNGSKILKEHCTFENVDNVVTIVPNEKAAMVNGVRID 571
DQ 503 -----KRREIEIQ--QMESRDEETLEKE--TYSSIQEQEVDI-----KTK 540

```

QY 572 PTLRSGVRIILGDPHFIRFNHPEEARAEQES-----LLRHSVTSNSOLG----- 617
 Db 541 LKLFESKQAVKABIHDLQEFIEHKERQLEQTQNELTRELKHLKIITIEPLEKKNKIM 600
 QY 618 -----SPAPGRHDTLSKAGSDADGSRSDS-----PLPHFRGKDS 653
 Db 601 NRSFDDDEEDHWLHPITRLENOQMKRPYSVAGYKRPILSOHARMSMIRPEYRAEN- 659
 QY 654 DMFYARREASAILGLDQKISHLTDELDALFDDVQKARVRGLVEDNEDSDQSFSF 711
 Db 660 -----IMLELDMPSRTRDYEPAISPKVQ---AALDAALQDEIDQVADASSF 705

RESULT 12
 ID K121_STRPU STANDARD; PRT; 742 AA.
 AC P46871;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE KINESIN-II 95 KDA SUBUNIT (KRP-85/95 95 KDA SUBUNIT).
 GN KRP95.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Egg;
 RX MEDLINE=94050179; PubMed=8232586;
 RA Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,
 RA Scholey J.M.;
 RT "Novel heterotrimeric kinesin-related protein purified from sea
 urchin eggs".
 RT Nature 366:268-270(1993).
 RL CC -1- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
 CC OF 95 AND 85 KDA.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC II SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U00996; AAA87393.1; --
 DR HSSP; P56536; 2KIN.
 DR InterPro; IPR001752; --
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.
 FT DOMAIN 1 337 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
 FT DOMAIN 338 613 COILED COIL (BY SIMILARITY).
 FT DOMAIN 614 742 GLOBULAR (BY SIMILARITY).
 FT NP_BIND 95 102 ATP (POTENTIAL).
 SO SEQUENCE 742 AA; 84202 MW; 47C40A367BAA77B5 CRC64;

Query Match 20.5%; Score 827.5; DB 1; Length 742;
 Best Local Similarity 31.5%; Pred. No. 3.5e-36;
 Matches 250; Conservative 122; Mismatches 286; Indels 135; Gaps 25;

QY 7 IKVVVRPPNAREIDRGAKCIVRNEGNOTI--LTPPGAEKARKSGKTMGPKAF 64
 Db 9 VKVWRCRPMNSREISQGHKRIVEMDNKRGVLVETNPKPGPEPKNS-----FTF 58

QY 65 DRSY-WSPDKNAPNARYAQEDLFQDLGVPLLDNAFKYNNCIFAYGQTGSGKSYSMGYGK 123
 Db 59 DTVYDWM-----SKQIDLYDETFSLVESVLFQFNGTIFAYGQTGKTFTMEGVRS 110
 QY 124 E---HGVIPIRQDMFRINELQDKNLJCTVEVSYLEINERNVRLANPSTKGNLKVRE 180
 Db 111 NPELRGVIPNSFEHFTIARTQNOQFL--VRASYLEIQVEIRDLAKOQKRELDK 167
 QY 181 HPSTGPVVEDLAKLVVRFQEIENLMDENKARTVAATNMNYSRSSHAFVTLTLITQKH 240
 Db 168 RPDGVVYKDLSSFTVKSEIKIEHVTGNNRVSSTNMNHSRSHAFITIT---E 223
 QY 241 DEETKMDTE---KVAKISLVLGSRATSTGATCARLKEGAEINRSLSTLGRVTAALAD 297
 Db 224 CSELGVGDGENHFRVGLNLDLAGSERQAKTGCDRLKEATKINLSLALGNVISALVD 283
 QY 298 MSSGQKKKNOLVPYRDSVLTWLLKDSLGNSTAMIAISPADINFEETLSTLYADSAK 357
 Db 284 -----GKSHPIPYRDSKLTLLQDSLGNNAKTVMANMGPSYNFDETTITLYANRAK 337
 QY 358 RIKNHAVYNDPNARMIRELKEELAOLRSKLQSSG-GGGGAGGSGGPGVEESYPPDTPLE 416
 Db 338 NIKNPKINEDPKDALLREFQOEISRLKQALDKGSPDKRKKKRRKPGEGGDDIDE 397
 QY 417 KQIVSIQOPATVYKMSKAEIVEQLNQSEKLYRDNLQNTWEEKLAKTEEI-----H 466
 Db 398 TEEGDEMDDEEYKESQOKLEE---EKEKIMANOSMTAEKOKLLSEYOKRQGEIKKEH 454
 QY 467 KEREAALEELGISIEKFGVGPY-----HSKEMPHLVNLSDDPLLAECVLYNINPKQOTRVGN 522
 Db 455 QOKEMLEGTKAMESKLLVGKGSIVDHTNQOKKIE-EDRLLAEE-----EKN 501
 QY 523 VNQDTQAEIRLNGSKILKEHCTENVDNVTVIPNEKAAVNVGVNRIDKPTPLRSQYRII 582
 Db 502 RERDMERKLEQDDKTVIEGTSSSQOQVE-----VTKKLAKLFAKLQSY 548
 QY 583 LGDFHIFRHNPEEAR-AERQEQSLLRHSVTSNSQLSGPAPGRHDT--LSKAGSDADGDS 639
 Db 549 KSDIOLQDEHARERQELEQTQNELIRELKLKKNADNFIPVEERTKITRAVFEDETEE 608
 QY 640 RSDSPLPHFRGKD-----SDMFYAR-----REASATILGLDQKIS 674
 Db 609 WLLTFLAKAEGFSQAKRPVSAGVGNRRPIAD--YARMAAQMGNGNPRYKAENLSVDLDM 666
 QY 675 HLT--DELDALFDDVQKARVRGLVEDNEDSDQSFSFPV-----RDKYMSNGTID 724
 Db 667 NRTTRDYEGFSVAPRVQ---AALDAALQDEDDLDLEQVQEVFKAKTKLKKDKVRSK---- 719
 QY 725 NFSLDTAITMPGT 737
 Db 720 ---HKAVAKPGS 728

RESULT 13
 FL10_CHLRE STANDARD; PRT; 786 AA.
 ID FL10_CHLRE
 AC P46869;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE KINESIN-LIKE PROTEIN FLA10 (KHPI PROTEIN).
 GN FLA10.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=137;
 RX MEDLINE=94299638; PubMed=8027176;
 RA Walther Z., Vashishtha M., Hall J.L.;
 RT "The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous

QY 169 NPSTKG-----NLVREHPSTGPPYEDLAKLVRSFOEINENDEGNKARTVAATNM 220
 Db 189 GKXKGVGGGEVYVDVREHPSRGVFLGQRLVEGSLDDVRLVIEGNGVRRHTASTKM 248
 QY 221 NETSSRSHAVFTPL-----TQKHDEETKMDTEKVAKISLVDLAGSERATSTGATCARLK 276
 Db 249 NDRSSRSHAIIMLLREERTMTTKSETIRTAGKSSRMMLVPLAGSERVAQSQVEGQPFK 308
 QY 277 EGAEINSLSTLGRVIAAALDMS--GKOKNOLVPYRDSVLTLTKSLGNSMTAMTAA 335
 Db 309 EATHINLSITLGRVIDVLADNATKGAQYSVAPRDSKLFILKSLGNSKTFMTAT 368
 QY 336 ISPADINFEETSLTRYASAKRIKHAHVNEPDNARMTRKELKLAQLSKSLQSSGGG 395
 Db 369 VSPSALNYEETSLTRYASARDIVNVAQVNEOPRARRIRELEQMEDMQAM----- 421
 QY 396 GGAGSGGPGVEE-----SYPPDTPLEK 417
 Db 422 --AGGDPAYVELKKLALLESEAQKRAADLQALEREREHNOVQERLLRATEAKSELES 479
 QY 418 QIVSTQ-----PDATVKKMSAEIVQOL 441
 Db 480 RAAALQECMTTRQADKQALNLRLKEEQARKERELLKEMAKKDAALSKVRRKDAETA 539
 QY 442 NOSEKLYRDLNOTWEEKLAKTEIEHKEREALBELGISIEKFGVPHSKEMPHLVNLS 501
 Db 540 SEREKL-----ESTVAQLEREQREVALDAL-----QTHQKQLQEALESSE 581
 QY 502 -----DPLAECILVYNKPGOTRVGNVNDQTQAEIRLNGSKILKECHTFENVVNTIV 555
 Db 582 RTAERDQLQO--LTELQSERITQLSVYVTDRELTR-DLQRIQYEGETELARQALCA 638
 QY 556 PNB-----KAAMVNGVRIDKPTLRSGYRIILGDFHIFRPNHPEEAERQBSILRHS 610
 Db 639 AQMEARYHAAYFLQTLLELATEDEALR-----ERALAEDEAAAAE-- 682
 QY 611 VTNQSLGSPAPGRDRTLKSGADGDSRSDSPLPFRPKDSDFWAFREAAASATILGID 670
 Db 683 -----LDAANSTSQN-----ARSACERLTSLE 705
 QY 671 QKISHLTDELALFDVQKARAVRGLVEDNEDSDSQSPFVRKTYMSNGTIDNFSLOT 730
 Db 706 QOLRE--SEERAELASOLEATAAAKSAEQDRENTATLEQOLRES-----EAPAAEL 757
 QY 731 AITMPTGPRSDGDALFFGDKKSKQDASNY-----DVEELRQOQAMBEALAKTAK 781
 Db 758 ASOLEATAAA-----KMSAQDRENTATLEQLRDSEREAELASQLESSTAAK 807
 RESULT 15
 KF3A_MOUSE STANDARD; PRT; 701 AA.
 ID KF3A_MOUSE AC P28741;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE KINESIN-LIKE PROTEIN KIF3A (MICROTUBULE PLUS END-DIRECTED KINESIN MOTOR 3A)
 GN KIF3A OR KIF3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain.
 RX MEDLINE=93077686; PubMed=1447303;
 RT Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M., Hirokawa N.;
 RA "Kinesin family in murine central nervous system.";
 RL J. Cell Biol. 119:1287-1296(1992).
 CC -/- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
 CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
 CC ACTIVITY IN VITRO.

CC -/- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
 CC -/- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
 CC TISSUE (MAINLY IN THE CEREBELLAR GRANULAR LAYER) WITHIN A SINGLE
 CC TYPE OF NEURONAL CELL.
 CC -/- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC II SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: D12645; BAA02166.1; -.
 DR PIR: B44259; B44259.
 DR HSP: P56536; 2KIN.
 DR MGD: MGI:107689; Kif3a.
 DR InterPro: IPR001752; -.
 DR Pfam: PF00225; kinesin; 1.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE: PS00467; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neutone.
 FT DOMAIN 1 350 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
 FT DOMAIN 351 586 COILED COIL (BY SIMILARITY).
 FT DOMAIN 587 701 GLOBULAR.
 FT NP_BIND 100 107 ATP (BY SIMILARITY).
 FT DOMAIN 442 445 POLY-GLU.
 FT DOMAIN 509 512 POLY-ARG.
 SQ SEQUENCE 701 AA; 80167 MW; 2405872DF2D85A29 CRC64;
 Query Match 19.9%; Score 800; DB 1; Length 701;
 Best Local Similarity 31.8%; Pred. No. 8.7e-35;
 Matches 231; Conservative 110; Mismatches 231; Indels 154; Gaps 25;
 QY 6 NIKVVVRPFNAREID---RGAKCIVRMENQOTLITPPGAEKARKSGKTMDGPKAF 62
 Db 14 NVKVVRCPLNEREKSMCYQAVSDVMRGTTV-----HKTDSSN----EPKPTF 61
 QY 63 AFRSYWSFDKAPYARQEDLFQDLGVPLDPAKFNKNCIFAYGOTGSGKSYMMGYG 122
 Db 62 TFDTVF-----GPE--SKOLDVYNLTARPIIDSVLEGVNGTIFAYGOTGKTPTMEGVR 114
 QY 123 KE---HGVIPRICQDMFRINELQKDKNLTCTVEYSYLEIYNERVDRDLNPNSTKGNLKV 179
 Db 115 AVPGIRGVIPNSFAHIFGHIAKAEQDTRF--LVRVSYLEYNEEVDRDLGDKDQTORLEVK 172
 QY 180 EHPSTGYPVEDLAKLVRSFOEINENDEGNKARTVAATNMNETSSRSHAVFTLT--TQ 237
 Db 173 ERPDGVYIKDLSAVVNNADDDMRITMLGHNRSVGATNNHSSRSHAITTIECSE 232
 QY 238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATCARLKECAEINBSLSTLCRVIAALAD 297
 Db 233 KVDENMHV---NMGLHLVLDAGSEKQAKTQGLKATKINLSLSTLGNVISALVD 289
 QY 298 MSSGQKQKQKOLVPRDSVLTWLLKDSLGNSMTAMIAAISPADINFEETSLTRYASAK 357
 Db 290 -----GKSTHPYPRNSKLTLLQDLSLGGNSKTMKCANIGPADYNYDETISTRLYANRAK 343
 QY 358 RIKHAVVNDPNARMIRELKEELAQLRKSLQSSG--GGGAGSGSGGPVEE----- 407
 Db 344 NIKKARINEDPKDALLQFOKELEELKKLEEGEVSGSDISGSEDEGEELGEGEK 403
 QY 408 -----SYPPDTPLEKQ-----IVS 421
 Db 404 RKRRDQAGKKVSPDKMVEAKIDERKALETKLDMEEPERNKARAELERRERKDLKA 463
 QY 422 IQQPDATVKKMSKAE-----IVBOLNQSEKLYRDLNOTWEEKLAKTEIEHKEREAA 472
 Db 464 QOEHSQSLLEKLSALEKKVYGVVDLLAKAEQEKLLSESNWEELEERRRRAEQLEKEEK 523

```
Qy 473 -----LEELGISTEKGVGYPYHSEMPHLYNLSDPDLIAECLVYNIKPGOTRVGNVNOD 526
      :||| |::: | |::: | |::: | |::: | |::: | |::: | |::: | |::: | |::: |
Db 524 EQERLDIEEKYTSLOEERAG--KTKKLKVVWM-----LMAAKSEMADLQOE 568
Qy 527 TQAEIR--LNGSKILKEHCTFEN--VDNVV-----TIVPN-----EKAAYMV 564
      :||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 569 HQREIEGLENIROLSRELRLQMLLIIDNFIPQDYQEMIENYVHWNEDIGEWQLKCVATG 628
Qy 565 NGVRIDKPT---RLRSGYRIILGDFHIFRFNHPEEARAEROEQSLLRHSVTNSQLGSPAP 621
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 629 NNMRRKQTPVPDKKERDPEVDLS--HVV-LAYTEESL--ROSLMKLERPRTSKGKARPKT 683
Qy 622 GRHDRT 627
      ||| | |
Db 684 GRRKRS 689
```

Search completed: April 25, 2001, 10:18:28
Job time: 294 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:12:54 ; Search time 77.5 Seconds
(without alignments)
1185.690 Million cell updates/sec

Title: US-09-235-416-1

Perfect score: 4030

Sequence: 1 MSGGNKVVVRVPPFNARE.....ELRQQAOMEALAKTAKEF 784

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-nhc:*
- 8: sp-organelle:*
- 9: sp-phase:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-unclassified:*
- 13: sp-vertebrate:*
- 14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1666.5	41.4	1770	11 Q92119	Q92119 mus musculus
2	1665	41.3	1103	4 Q75186	Q75186 homo sapien
3	1660.5	41.2	1816	11 Q9R0B4	Q9R0B4 mus musculus
4	1652.5	41.0	1816	11 Q9WVE5	Q9WVE5 mus musculus
5	1635	40.6	1773	5 Q9V776	Q9V776 drosophila
6	1630.5	40.5	1671	5 Q9NBL1	Q9NBL1 drosophila
7	1617	40.1	1097	11 Q35787	Q35787 rattus norv
8	1590.5	39.5	1584	5 Q18778	Q18778 caenorhabdi
9	1525.5	38.6	689	11 Q88658	Q88658 rattus norv
10	1525.5	37.9	2205	5 Q9NGQ2	Q9NGQ2 dictyostel
11	1412	35.0	1826	4 Q9N078	Q9Nq48 homo sapien
12	1396.5	34.7	1921	5 Q01349	Q01349 drosophila
13	1306.5	34.7	1921	5 Q9V7C9	Q9V7C9 drosophila
14	1284.5	31.9	1576	5 Q20888	Q20888 caenorhabdi
15	1174.5	29.1	1648	4 Q10058	Q10058 homo sapien
16	1147	28.5	928	5 Q99997	Q99997 caenorhabdi
17	1128	28.0	1174	5 Q9VB25	Q9VB25 drosophila
18	1035	25.7	1121	5 Q9VIP4	Q9VIP4 drosophila
19	1035	25.7	1121	5 Q16866	Q16866 drosophila

20	1034	25.7	1121	5 Q18390	Q18390 drosophila
21	981.5	24.4	412	4 Q9UJ02	Q9UJ02 homo sapien
22	847.5	21.0	752	4 Q9NXN9	Q9NXN9 homo sapien
23	799	19.8	744	13 Q93478	Q93478 xenopus lae
24	796	19.8	2756	10 Q9LJ60	Q9LJ60 arabidopsis
25	781.5	19.4	677	5 Q9VRK9	Q9VRK9 drosophila
26	775	19.2	1229	10 Q9LXV6	Q9LXV6 arabidopsis
27	775	19.2	1232	4 Q9NY24	Q9NY24 homo sapien
28	763	18.9	1099	5 Q9U142	Q9U142 leishmania
29	760	18.9	797	11 Q88657	Q88657 rattus norv
30	755.5	18.7	1225	13 Q90640	Q90640 gallus gall
31	750.5	18.6	784	5 Q9VTN8	Q9VTN8 drosophila
32	748	18.6	735	5 Q9U0D5	Q9U0D5 tetrahymena
33	744	18.6	1226	13 Q91784	Q91784 xenopus lae
34	743.5	18.4	697	5 Q9U921	Q9U921 tetrahymena
35	735.5	18.3	1128	5 Q9U179	Q9U179 leishmania
36	728.5	18.1	2158	10 Q9LUT5	Q9LUT5 arabidopsis
37	722	17.9	929	3 P78718	P78718 nectria hae
38	709	17.6	1292	10 Q9LUD0	Q9LUD0 arabidopsis
39	701	17.4	1130	5 Q19633	Q19633 caenorhabdi
40	692.5	17.2	2273	5 Q9U141	Q9U141 leishmania
41	690	17.1	402	4 Q9Y6V4	Q9Y6V4 homo sapien
42	681	16.9	968	3 P87199	P87199 ustilago ma
43	679	16.8	1212	5 Q9X229	Q9X229 drosophila
44	678	16.8	1075	10 Q9SNE3	Q9SNE3 arabidopsis
45	672.5	16.7	1668	11 Q9QXL1	Q9QXL1 mus musculu

ALIGNMENTS

```

RESULT 1
Q92119 ID Q92119 PRELIMINARY; PRT; 1770 AA.
AC Q92119;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE KIF1B-BETA.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR MOUSE;
RA Nakagawa T., Hirokawa N.;
RT Identification and characterization of a new kinesin superfamily
RL KIF1B-beta.
RL Submitted (JEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023656; BAA75243.1; -.
DR HSSP; P33176; IBG2.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1770 AA; 198850 MW; E444228501672E19 CRC64;

```

Query Match 41.4%; Score 1666.5; DB 11; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.7e-90;
Matches 353; Conservative 125; Mismatches 174; Indels 95; Gaps 15;

Qy 4 GGNKVVVRVPPFNAREIDRGAKCIVRMENQITLTPPGAEKARKSGKTMIDGPKAPA 63

```
Db 3 GASVAVVRVPFNSRETSKESKCIQMOGNSTSIINPKNPK-----APKSFS 51
Qy 64 FDRSYWSF-DKNAPYARQEDLFDQLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMGVG 122
Db 52 FDISYWSSTSPEDPCFASQNRINDIGKEMLLHAFEGYNNCIFAYGQTGAGKSYTMGQK 111
Qy 123 KEH--GVIPRICQDMFRRIEQLQDKNLTCTVEVSYLEIYNERVDRLLNPNSTKGNLKYRE 180
Db 112 ESQAGIIPOLCEELFEKIND-NCEEMSYSVEVSMEIYCERVDLLNPNKNGNLRYRE 170
Qy 181 HPSTGPPYVEDLAKLVVRFSQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
Db 171 HPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQKKQ 230
Qy 241 DEETKMDTEKYAKISLVDLAGSERATSGATGARGKEAGINRSLSLTGRVTAALADMS 300
Db 231 DPETNLSTKYSKISLVDLAGSERADSGAGTRLUKEGANINKSLTTIGKVISALAEYSK 290
Qy 301 GKQKNQLVPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLRVYADSAKRK 360
Db 291 -KKKTDIPYRDSVLTWLLRNLGGNSRTAWAALSADINDETSLSTLRVYADRAKQK 349
Qy 361 NHAVNEDPNARMIRELKEELAQLRSKLSQSGGGGGAGGSGPPYEDTTPLEKQ-- 418
Db 350 CNAVINEDPNAKLVRELKEEVTRLKDLRAQGLGDIIDTSMGSLT--SSPSSCSLSNSQV 407
Qy 419 ---IVSIOQDATYKMSKAETVEOLNOSEKLYRDLNOTWEEKLAKTEIHKERAALAE 475
Db 408 LTVSTSIQ--ERIMSTPGGEAIERLKESEKIIAINETWEEKLRKTEAIRMERALLAE 465
Qy 476 LGISTEK--GFVGPYHSEMPHLVNLSDPLLAELCLVNIKPGOPRVGNVODTQAEIRL 533
Db 466 MGVALREGDGLGVSPKPKTHLVNLNEDPLMSECLLYIKDGITRVGOAERQDIDL 525
Qy 534 NGSKILKECHTFENV-----DNVVTIVPNEKAAMVNGVRIDKPTRLSRGYRIILGDFHI 588
Db 526 SGAIKEEHLFRSRSNTGEVITVLEPCERSETVYNGKRVAPVOLRSNGRIINGKNHV 585
Qy 589 FRFNHPEARARERQOSILLRHSVNSQLGSPAPGHRDRTLKAGSDADGDSRSDPLPHF 648
Db 586 FRFNHPEARAREK-----TPSAET 606
Qy 649 RGKSDWFYARREASAILGLDOK-----LSHLTDDLDAL-----FDDVOKA 691
Db 607 PSEPVDYTFARLELEK--QGIDMKQEMKRLQEMILYKKEKEADLLLEQORLDYESKL 665
Qy 692 RAVRRGL-----VEDNEDSDSQSFP 712
Db 666 QALORQVETRSLAETTEEEEREVP 692
```

RESULT 2

```
O75186 ID O75186 PRELIMINARY; PRD; 1103 AA.
AC O75186;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE KIAA0706 PROTEIN.
GN KIAA0706.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT *Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
```

```
RT code for large proteins in vitro.;
RL DNA Res. 5:169-176(1998).
DR EMBL; AB014606; BAA31681.1; -.
DR HSSP; P17119; 3KAR.
DR INTERPRO; IPR000253; -.
DR INTRPRO; IPR001752; -.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1103 AA; 122946 MW; 5FDEEB8C91B3C46B CRC64;
```

Query Match 41.3%; Score 1665; DB 4; Length 1103;
Best Local Similarity 47.5%; Pred. No. 1e-90;
Matches 354; Conservative 120; Mismatches 158; Indels 114; Gaps 16;

```
Qy 4 CGNKKVVRVPFNAREIDRGAKCIVRMEGNQITLTPPPGAEKARKSGKTINDGPKAFA 63
Db 3 GASVAVVRVPFNARETSQDAKCVSMQGNNTSIINP-----KOSKDAKPSFT 51
Qy 64 FDRSYWSFDKNA-PNVARQEDLFDQLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMGVG 122
Db 52 FDISYWSSTSPEDPCFASQOQYRDIGEMLLHAFEGYNNCIFAYGQTGAGKSYTMGRO 111
Qy 123 K--EHGVIPRICQDMFRRIEQLQDKNLTCTVEVSYLEIYNERVDRLLNPNSTKGNLKYRE 180
Db 112 EPGQOIGVQLCEDLFSRVSENO-SAQLSYSVEVSMEIYCERVDLLNPNKSRGSLRYRE 170
Qy 181 HPSTGPPYVEDLAKLVVRFSQEIENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
Db 171 HPLLGPYVEDLSKLAVTSYTDIADLMDGKNKARTVAATNMNETSSRSHAVFTIVFTQCH 230
Qy 241 DEETKMDTEKYAKISLVDLAGSERATSGATGARGKEAGINRSLSLTGRVTAALADMS 300
Db 231 DQTLGDSKYSKISLVDLAGSERADSGAGMRUKKEGANINKSLTTIGKVISALADMS 290
Qy 301 GKQKNQLVPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLRVYADSAKRK 360
Db 291 -KKRKSDFIPYRDSVLTWLLKENLGGNSRTAMIAALSADINYEETLSTLRVYADRTKQIR 349
Qy 361 NHAVNEDPNARMIRELKEELAQLRSKLSQSGGGGGAGG----- 400
Db 350 CNAVINEDPNARLIRELQEEVARLRELLMAQGLSASALEGLKTEEGSVRGALPAYSSPPA 409
Qy 401 -----SGPVEESYPPPTPLEKQIVSIQPDATVKKMSKAEIVBQLNOSEKLYRDL 451
Db 410 PVSPTTHNGELEPSPNWT--ESQI-----GPEAMERLOETEKIIAEL 454
Qy 452 NOTWEEKLAKTEIHKERAALAEELIGISIEK--GFVGPYHSEMPHLVNLSDPLLAELCL 509
Db 455 NETWEEKLRKTEALRMEREAALLAEMGVAVREDGGTVGVSPKPKTHLVNLNEDPLMSECL 514
Qy 510 VYNIKPGQTRVGNVNDTQAEIRLNGSKILKECHTFENV-----DNVVTIVPNEKAAMV 564
Db 515 LYHIKDGTVRVGQVDM-----IKLTQPIREQHCLFRSIPQPDGGEVVVTVLEPCEGETYV 570
Qy 565 NGVRIDKPTRLSRGYRIILGDFHIFRNHPEARARERQOSILLRHSVNSQLGSPAPGRH 624
Db 571 NGKLVTPELVKSGNRIVMGKNHVRFNHPBOARLER-----GVPPP----- 614
Qy 625 DRTLKAGSDADGDSRSDPLPHFRGKSDWFFYARREASAILGLDQKISHLTDDLDAL 684
Db 615 -----PGPPSEPDWNPFAKLELEQ--QGIDIKLE--MEKRIQLD 650
Qy 685 FDDVOKARAVRGLVEDNE---DSDS 707
Db 651 ENQYRKEKEEADLLLEQORLYADSDS 676
```

RESULT 3


```
Q9R0B4
ID Q9R0B4 PRELIMINARY; PRT: 1816 AA.
AC Q9R0B4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR OUTBRED; TISSUE=BRAIN;
RA Gong T.L., Winnicki R.S., Kohman D.C., Lomax M.I.;
RT "A novel kinesin of the UNC-104/KIF1 subfamily encoded by the kif1b
gene.";
KL Gene 239:117-127(1999).
DR EMBL; AF090190; AAF06718.1; -.
DR HSSP; P33176; 1BG2.
DR MGD; MGI:108426; Kif1b.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PS00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1816 AA; 204079 MW; E316EC295138E5DE CRC64;

Query Match 41.2%; Score 1660.5; DB 11; Length 1816;
Best Local Similarity 45.1%; Pred. No. 4e-90;
Matches 357; Conservative 127; Mismatches 170; Indels 137; Gaps 17;

QY 4 GGNKVVVVRFPNAREIDRGAKCIVRMGNOTILTPPGCAEAKARKSGKTIMDGPKAFA 63
DB 3 GASVKVAVRVRFPNSRETSKESKCIIOGNSSTIINPKNPKE-----APKSFS 51

QY 64 FDRSYWSF-DKNAPNYARQEDLFODLGVPLLDNAFGKNCIFAYGQSGKSYSMWGYG 122
DB 52 FDSYWSHTSPEDPCFASQNRVYNDIGKMLLHAFEGYNVCIFAYGQSGACKSYTMWGKQ 111

QY 123 KEH--GVIPRICDMERRINELQKDKNLCTVEYSYLEINRVRDILLNPSTKGNLKYRE 180
DB 112 EESQAGIIPOLCEELPEKIND-NCNEEMSYSEVSYMEIYICERVVDLLNPKNGNLVRRE 170

QY 181 HPTSGPYVEDLAKLVRSFOEINLMDGKNKARTVAATNNNETSSRSKSHAVFTLTQKWH 240
DB 171 HPLGLPYVEDLSKLVASTYTDIADLADGAKNKARTVAATNNNETSSRSKSHAVFTIVTQKKQ 230

QY 241 DEETKMDTEKVAKTSLVDLAGSRATSTGATGARKKEGAFINRSLSLGRVIALADAM-- 298
DB 231 DPETNLSTEKYSKLSVLDLAGSRADSTGATGARKKEGANINKSLTTLGKVISALAEYDN 290

QY 299 ---SSQKQKNOLVPYRDSVLTWLDKSLGNSWTAMIAISPAINPEETPLSTLYRADS 355
DB 291 CTSKSKKKKKKDFIPYRDSVLTWLLRNLGGNSRTAMVAALSPADINDYDTLSTLYRADR 350

QY 356 AKRIKKNHAVNEDNARMIREKEELAQRSLQSSGGG-----GGAGG---- 400
DB 351 AKQIKNAVINEDNPAKLVRELKEEYTRKDLPLRAQGLGDIIDPLIDYSSGGSKYLK 410

QY 401 -----SGGVPEESYDPDTPLEKQ-----TVSQQPDATVKK 431
DB 411 DFQNNKHYLLASENQRPGNFSTASMGSLTSS-PSSCSLSNQVGLTSTVTSIQ--ERIMST 467

QY 432 MSKAEIVELQNSKLYRDLNQTWEEKLAKTEETHKEREAALEELGISIEK--GFVGPHY 489
```

```

Db 468 PGGEAERLKESEKIIAELNETWEEKLRKTEAIRMERALLAEMGVAIREDGGTIGVFS 527
QY 490 SKEMPHLVNSDDPLLAECIVYNIKPGQTRGVNVNODTQAEIRLNGSKILKEHCTFENV- 548
DB 528 PKKTHLVNLNEDPLASECLLYIKDGITRVQADARRQDVLVSGAHIKEEHCILRSER 587
QY 549 ----DNVVTIVPNEKAAMVNGVRIDKPTRLRSYRIILGDFHIFRNFHPPEARARQEQ 604
DB 588 SNTGEVIITLPECSERSETYVNGKRVAPVQLRSNGRIIMGNHVFREHNHPEQARAREK- 646
QY 605 SLLRHSVTNSQLGSPAPGRHRTLSKAGSDADGSDSDPLPHFRGKSDWVFYARREAS 664
DB 647 -----TPSNETPSEPVDWTFQAQELLE 668
QY 665 AILGLDQK-----TSHLTDDDELDALE-----FDVOKARAVRRCL-----VED 701
DB 669 K-QGIDMKQEMEKRLQEMELLYKKKEEADLLLEQQRLDYESKLQALQROVETRSLAAET 727
QY 702 NEDSDSQSSFP 712
DB 728 TEEEEEVEVP 738

RESULT 4
Q9WVE5 PRELIMINARY; PRT: 1816 AA.
AC Q9WVE5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KIF1B MAJOR ISOFORM.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=BRAIN;
RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,
RA Perry V.H., Coleman M.P.;
RT "The major brain isoform of Kif1b lacks the putative mitochondria-
binding domain.";
RL Mamm. Genome 10:0-0(1999).
DR EMBL; AF131865; AAD39438.1; -.
DR HSSP; P31176; 1BG2.
DR MGD; MGI:108426; Kif1b.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1816 AA; 204153 MW; 95CB196A9DB4895A CRC64;

Query Match 41.0%; Score 1652.5; DB 11; Length 1816;
Best Local Similarity 45.0%; Pred. No. 1.2e-89;
Matches 356; Conservative 127; Mismatches 171; Indels 137; Gaps 17;

QY 4 GGNKVVVVRFPNAREIDRGAKCIVRMGNOTILTPPGAEAKARKSGKTIMDGPKAFA 63
DB 3 GASVKVAVRVRFPNSRETSKESKCIIOGNSSTIINPKNPKE-----APKSFS 51

QY 64 FDRSYWSF-DKNAPNYARQEDLFODLGVPLLDNAFGKNCIFAYGQSGKSYSMWGYG 122
DB 52 FDSYWSHTSPEDPCFASQNRVYNDIGKMLLHAFEGYNVCIFAYGQSGACKSYTMWGKQ 111

QY 123 KEH--GVIPRICDMERRINELQKDKNLCTVEYSYLEINRVRDILLNPSTKGNLKYRE 180
DB 112 EESQAGIIPOLCEELPEKIND-NCNEEMSYSEVSYMEIYICERVVDLLNPKNGNLVRRE 170

QY 181 HPTSGPYVEDLAKLVRSFOEINLMDGKNKARTVAATNNNETSSRSKSHAVFTLTQKWH 240
DB 171 HPLGLPYVEDLSKLVASTYTDIADLADGAKNKARTVAATNNNETSSRSKSHAVFTIVTQKKQ 230

QY 241 DEETKMDTEKVAKTSLVDLAGSRATSTGATGARKKEGAFINRSLSLGRVIALADAM-- 298
DB 231 DPETNLSTEKYSKLSVLDLAGSRADSTGATGARKKEGANINKSLTTLGKVISALAEYDN 290

QY 299 ---SSQKQKNOLVPYRDSVLTWLDKSLGNSWTAMIAISPAINPEETPLSTLYRADS 355
DB 291 CTSKSKKKKKKDFIPYRDSVLTWLLRNLGGNSRTAMVAALSPADINDYDTLSTLYRADR 350

QY 356 AKRIKKNHAVNEDNARMIREKEELAQRSLQSSGGG-----GGAGG---- 400
DB 351 AKQIKNAVINEDNPAKLVRELKEEYTRKDLPLRAQGLGDIIDPLIDYSSGGSKYLK 410

QY 401 -----SGGVPEESYDPDTPLEKQ-----TVSQQPDATVKK 431
DB 411 DFQNNKHYLLASENQRPGNFSTASMGSLTSS-PSSCSLSNQVGLTSTVTSIQ--ERIMST 467

QY 432 MSKAEIVELQNSKLYRDLNQTWEEKLAKTEETHKEREAALEELGISIEK--GFVGPHY 489
```

Db 52 FDYSWSHTSPDPFCFASQNRVYNDIGKEMLLHAFEGYNVCIAYGQTGACKSYTMMGKQ 111
Qy 123 KBHG--VTPRICODMFRRLNELQKDNLTCTVEVSYLETYNERVRLDLPSTKGNLKYRE 180
Db 112 ESQAVIIPOLCEELFEKIND-NCHEMSYSVEVSINIELYCERVRLDLPNKNKGNLRYRE 170
Qy 181 HPSTGPGYVEDLAKLVYRSFOETENIMDEGNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
Db 171 HPLLPGYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVTQKKQ 230
Qy 241 DEETWMTKEKAKISLVDLAGSERATSTGATGARKLKEGAENRSLSTLGRVTAALADM-- 298
Db 231 DPETLSTKEKISKISLVDLAGSERADSTGAKGTRUKEGANINKSULTTGLKVISALAEYDN 290
Qy 299 ---SSGKOKKNOVPYRSVLTWLKDSLGNSMTAMTAAISPADINEETLSTLYADS 355
Db 291 CTSKSKKKKKKTFIPYRSVLTWLKRENLGGNSRTAMVAALSPADINDETULTLYADR 350
Qy 356 AKRIKNHAVNEDPNARMTRELKEELAQRSLKQSGGGG-----CGAGG---- 400
Db 351 AKQIKCNAVINEDPNAKVLRELKEVTRLKOLLRAQGLGDIIDPLDDYSGSGKYLK 410
Qy 401 -----SGGPVEESYPPDTPLEKQ-----IVSQOQDATVKK 431
Db 411 DFQNNKHRYLLASENQRFCNFSTAGMSLTSS-PSSCSLNSQVGLTSVTSIQ--BRIMST 467
Qy 432 MSKAEIVQOLNOSKLYRDLNOTWEKLAETIEHKREFAALEELGISTEK--GFVGPYH 489
Db 468 PGGEAERKESEKIIAELNETWEKLRKTAIRMEREAALAEVRLDGGTGLGVFS 527
Qy 490 SKEMPHLVNSDDPLLAECVLVYNIKPGQTRVGNVQDQAEIRLNGSKILKHCXFFENV- 548
Db 528 PKKTHLVNLNEDPLMSECLLYIKDGTIVRGQADAEERRQDVLVLSAHIKEHCLFRSER 587
Qy 549 ----DNVTVIVPNEKAAVWVNGVRDKPRLRSYGRILGDPIHFRFNHPEARAEQEQ 604
Db 588 SNTGEVITVLEPCSETSYVNGKVAHPVQLRSGNRRIIMGKKNHVFNFHPEQARAEK- 646
Qy 605 SLRHSVNTSOLGSPAGRHDRTLKAGSDADGDSRSDPLPHFRKSDWFYARREAS 664
Db 647 -----TPSAETPSEPVDMTFAQRELLE 668
Qy 665 AILGLDQK-----ISHLTDELDAL-----FDDVQKARAVRRL-----VED 701
Db 669 K-QGIDMKQEMKRLQEMELLYKKEKEADLLEQRLDYESKLQALQOVETRSLAAET 727
Qy 702 NEDSDQSQSPF 712
Db 728 TEEEEEVEVP 738

RESULT 5
Q9V7T6 ID Q9V7T6 PRELIMINARY; PRT: 1773 AA.
AC 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE CG8566 PROTEIN.
OS Drosophila melanogaster (Fruit fly).
GN Drosophila
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fester C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhao G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003805; AAF57957.1; -.
DR HSSP; P17119; 3KAR.
DR FLYBASE; FBgn0034155; CG8566.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1773 AA; 200755 MW; 335BE9CD5E4DB681 CRC64;

Query Match 40.6%; Score 1635; DB 5; Length 1773;
Best Local Similarity 44.5%; Pred. No. 1.3e-88;
Matches 367; Conservative 143; Mismatches 193; Indels 122; Gaps 20;
Qy 6 NIKVVVRPPNAREIDRGAKCIVRMENQITLTPPGAEEKARKSGKTIMDGKAFAPD 65
Db 32 SVKAVRVRRPNSREIARESKCIIEAGATTATNP-----KVPNTSDSVRRFND 83
Qy 66 RSYWSFDKNAPNRYARQEDLFDLGVPLLDNAPKGYNNCIFAYGQTGSGKSYSGMMGYGRE- 124
Db 84 YSNWSDHDDADFSTQSMVYKDIGEMLQHSFDGYNVCLIFAYGQTGAGKSYTMMGRQEQ 143
Qy 125 -HGVTIPRICQDMFRINELQKDNLTCTVEVSYLETYNERVRLDLPSTKGNLKYREHPS 183
Db 144 QEGILPMICKDLFTRIQDTETD-DLKYSVEVSMEIYCERVRLDLPNKNKGNLRYREHPL 202
Qy 184 TCPYVEDLAKLVYRSFOEINLMDGKARTVAATNMNETSSRSHAVFTLTITQKWHDEE 243
Db 203 LGPYVEDLSKLAVDYQDIHLIDEGNKARTVAATNMNETSSRSHAVFTITFQRRHDLIM 262
Qy 244 TKMDTEKVKAKISLVDLAGSERATSTGATGARKLKEGAENRSLSTLGRVTAALADMSSGKQ 303
Db 263 TNLTEKYSKISLVDLAGSERADSTGAKGTRUKEGANINKSULTTGLKVISALAEVASKK 322

```
QY 304 --KNQVLPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSLRVADSARKIKN 361
DB 323 NTKKAOFIPYRSALUTWLLRNGGSKMTAMIAISPADINDETLSLRVADRAKOIYC 382
QY 362 HAVYNEDPNARMIRELKEEALQRLSKOSSGGGGAGGGGPPVPESTPPPLEKQIVS 421
DB 383 KAVVNEDANAKLIRELKEEIQKRLDLKAEQ-----IEVQEGPGK-----VV 425
QY 422 IOOPDATVKKMSKAEI-----VEQLNOSEKLYRDLNOTWEEKLAKTE 463
DB 426 CEKRDANKBELTSTVTKSPKSRNRNGSTTEMAVDQQLQASEKLAELNETWEEKLAKTE 485
QY 464 EIKHREAALEELGIST-EKGF-VGPYHSEKMPHLVNLSDPDLAECVLYNKPQGTNRV 521
DB 486 EIRVQREAVFAGMGVAVKEDGITVGVSPKTPHLVNLNEDPNLSECLLYIKEGLTRLG 545
QY 522 NVNDQTOAEIRLNGSKILKECHTFENVDNVTVIPNEKAAVMVNGVRIDKPTLRSGYRI 581
DB 546 THEANVPDQIQSGSHILKECHTFENKNSVTLLPHKDAIIVNGRKLVEPVLTGSRV 605
QY 582 ILGDFHIFRNPHEPARAEQOSLLRHSVNSQLGSPAPGRHDTLSKAGSDADGDSRS 641
DB 606 ILGKNHVFRTPEQARE-----LRDKI-----ETENEAEVE---KT 641
QY 642 DSPLPHFRCKDSDFYARREASAILGLD-----OKISHLTDD-----ELDALFDDVQ 689
DB 642 DT-----QVDNFAQCELLEK-OGIDLKAEKMKRLDNLDEQYKREKLOADQOFEQR 693
QY 690 K---AR-AVRRLGVNEDSDSQSPFVRDKYMSNGTIDNFSLDATIMP----- 735
DB 694 KTEARIDALQOQVEOSMTMSYSSPDEHQEDVYTNPMYESCWTAREAGLAAMAF 753
QY 736 -----GTPRSDDGDAFLFGDKKSKODASNVDEELRQOQAO 772
DB 754 RKNRYHOFTSLRDDLWGNALFL-----KEANAISVELKKVQFO 792

RESULT 6
ID Q9NBL1 PRELIMINARY; PRT; 1671 AA.
AC Q9NBL1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN SUPERFAMILY MEMBER DUNC104.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossberger R., Saxton W.M., Dickson B.J.;
RT "Characterization of the Drosophila Unc104/KIF1A homolog, Dunc104.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247761; AAF74192.1; -.
SQ SEQUENCE 1671 AA; 189430 MW; BAC3F1D3022360AF CRC64;

Query Match 40.5%; Score 1630.5; DB 5; Length 1671;
Best Local Similarity 44.3%; Pred. No. 2.le-88;
Matches 364; Conservative 141; Mismatches 188; Indels 129; Gaps 20;

QY 6 NIKVVRVPPFNAREIDGKACIVRMGNOITLTPPGCAEAKRKSQGTIMDGKPAFED 65
DB 3 SVKAVVRPPFNAREIDGKACIVRMGNOITLTPPGCAEAKRKSQGTIMDGKPAFED 54
QY 66 RSYWFDKKNAPYARQEDLQDLGVPLDNLNFKYNNCIPAYGQSGSKYSMMGYCKE- 124
DB 55 YSYWHDHDADEFSTQSMYKDIGREMLQHSFDGYNVCIFAYGQSGSKYSMMGYCKE 114
QY 125 -HGVIPIQDMFRNELQKKNLTCTVETVSELYETINVERNROLLNPSTKGNLKVREHPS 183
```

```
DB 115 QGGIIPMTCKDLFFRQDTETD-DLKYSVEVSTWYICERVRLDLPNKNKGNLRVREHPL 173
QY 184 TGPYVEDLAKLVRSFOEIEENLMDGKNKARTVAATNMNFTSSRSYAVFTLTTLTKQKHDEE 243
DB 174 LGPYVEDLSKLVATDYQDHLIDEGNKARTVAATNMNFTSSRSYAVFTLTTLTKQKHDEE 233
QY 244 TKMDTEKVAKTSVLDSLAGSERATSGTAGRLKEGAENRSISTLGRVIAALADMSGSKO 303
DB 234 TNLITTEKYSKISLVDSLGERADSTCGAKTRELKEGANINKSLTTLGKVISALAEVSAKK 293
QY 304 ---KNQVLPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSLRVADSARKIKN 360
DB 294 KNTKKAOFIPYRSALUTWLLRNGGSKMTAMIAISPADINDETLSLRVADRAKOIYC 353
QY 361 NIAVYNEDPNARMIRELKEEALQRLSKOSSGGGGAGGGGPPVPESTPPPLEKQIV 420
DB 354 CKAVVNEDANAKLIRELKEEIQKRLDLKAEQ-----I 386
QY 421 STQPPD-----ATVKKMSKAE-----IVQLNQSEKLYRDLNOTWEEKLAKTEEIH 466
DB 387 EYQDEDELTKSTVIKSPKSRNRNGSTTEMAVDQQLQASEKLAELNETWEEKLAKTEEIR 446
QY 467 KEREAALEELGIST-EKGF-VGPYHSEKMPHLVNLSDPDLAECVLYNKPQGTNRVGNV 524
DB 447 VQREAVFAGMGVAVKEDGITVGVSPKTPHLVNLNEDPNLSECLLYIKEGLTRLGTHE 506
QY 525 QDQTOAEIRLNGSKILKECHTFENVDNVTVIPNEKAAVMVNGVRIDKPTLRSGYRIILG 584
DB 507 ANVPDQIQSGSHILKECHTFENKNSVTLLPHKDAIIVNGRKLVEPVLTGSKRVILG 566
QY 585 DFHIFRNPHEPARAEQOSLLRHSVNSQLGSPAPGRHDTLSKAGSDADGDSRSDSP 644
DB 567 KKNHVFRTPEQARE-----LRDKI-----ETENEAEVE---KTD- 601
QY 645 LPHFRCKDSDFYARREASAILGLD-----OKISHLTDD-----ELDALFDDVQK- 690
DB 602 -----QVDNFAQCELLEK-OGIDLKAEKMKRLDNLDEQYKREKLOADQOFEQRTY 654
QY 691 -AR-AVRRLGVNEDSDSQSPFVRDKYMSNGTIDNFSLDATIMP----- 735
DB 655 EARIDALQOQVEOSMTMSYSSPDEHQEDVYTNPMYESCWTAREAGLAAMAFRW 714
QY 736 -----GTPRSDDGDAFLFGDKKSKODASNVDEELRQOQAO 772
DB 715 RYHOFTSLRDDLWGNALFL-----KEANAISVELKKVQFO 750

RESULT 7
ID Q35787 PRELIMINARY; PRT; 1097 AA.
AC Q35787;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE KINESIN-RELATED PROTEIN.
GN KIF1D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR;
RA Rogers K.R., Griffin M., Brophy P.J.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ000696; CAA04248.1; -.
DR HSSP; P17119; 3KAR.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
```

DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1097 AA; 122333 MW; 8EF40B1C7579BA5B CRC64;

Query Match 40.1%; Score 1617; DB 11; Length 1097;
Best Local Similarity 45.9%; Pred. No. 7.3e-88;
Matches 357; Conservative 129; Mismatches 184; Indels 108; Gaps 20;

QY 4 GGNIKVVVRFPNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAPA 63
DB 3 GASVAVKVRFPNARETSQDAKCVVSMQGNITSILNP-----KOSRMPL---KA-S 50
QY 64 FDRSWSEDK-NAPNARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQSGSKSYMMGYG 122
DB 51 FDISYWSHTSVEDPQASQOQVYRDIQGEMLLHAEYGVNCFAYGQSGAGKSYMMGRQ 110
QY 123 K--EHGVIPRICQDMFRINELQDKNITCTVEVSYLEIYNERNVRLDLPSTKGNLKYRE 180
DB 111 EFGQOQIVPQCELDLFRVN-VNQSALQSYSEVSMEYTCERVRLDLPNPKSRGSLRVRE 169
QY 181 HPSTGPYVEDLAKLVRSFOELENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWH 240
DB 170 HPILGPYVQDLSKAVTSYADIADLMDCGNKARTVAATNMNETSSRSHAVFTIVETQRSH 229
QY 241 DEETKMDTEKAKISLVDLGASERATSGATCARLKEGAENRSLTGLRVIAALADMS 300
DB 230 DLTGLDSEKYSKISLVNLASERADSGARGMRLEKAGANINKSLTTLGKVISALADQS 289
QY 301 GKQKNQIIPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTRYADSAKRK 360
DB 290 -KKRKSDFIPYRDSVLTWLLKENLGNSRTAMIAISPADINYEETLSTRYADTKQIR 348
QY 361 NHAVNEDPNARMIRELEKEELAQLRKSKLOSSGGGGGAGG-----SGG----- 403
DB 349 CNAVINEDPNARLIRELEQEARLRELLMAQGLKVSASALGGLKVERGSPGVLPAASSPPA 408
QY 404 PVEESYPDTPLEKQIVSTQPDPAVKKMSAEIVQELNOSKLYRDLNOTWEEKLAKTE 463
DB 409 PASPSPP--PHNGELEFSFSAE-POIGPEAEERLQETEKITAEINWEEKLRKTE 465
QY 464 EHKEREAALEELGISIEKFGVGYHSEMPHLVNLSDPDLAELGVNIRKPGQTRVGNV 523
DB 466 ALRMEREALLAEMGSPGWRTVGFSPKTPHLVNLNEDPLMSECLLYHINDGVTRVGQV 525
QY 524 NQDTQAEIRLNGSKILKECHTFENV-----DNVTVIVNEKAAVWNGVRDKPRLRSG 578
DB 526 DVD-----IKLTGQFIREQHCLFRSIPQPDGEVMVTFEPCGEATYVNGKLVTPEVLKSG 581
QY 579 YRIILGDFHIFRNPHEPARERQBSLLRHSVTNSQLSGSPAPGRHRTLSKAGSDADGD 638
DB 582 NRIYGNKHVFRFNPEQARLER-----GVPPP----- 611
QY 639 SRSDSPLPHFGKSDNPFYAREASAILGLDQKISHLTDDDELDFDQVOKARVRRLG 698
DB 612 -----PGPPSEPDVNFNAQDKWLEQ-QGIDIKLE--MEKRLQDLENQYRKEEADLL 661
QY 699 VE-----DNEDSDQSS-----FPRYDKYMSNGTIDNLSLDTAITMPTGPRS 740
DB 662 LEQQRLYADSDGSDGSDRSCSESWRLISLRLDELPN-----TVQTIVRCGLPSS 713

RESULT 8

ID Q18778 PRELIMINARY; PRT; 1584 AA.
AC Q18778;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE C. ELEGANS UNC-104 KINESIN-LIKE PROTEIN (PIR:U01114).
GN UNC-104.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Cooper J., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.; III of C.
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Du Z.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50135; AA93453.1; -.
DR HSSP; P17119; 3KAR.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH: 1.
DR PFAM; PF00225; kinesin: 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1584 AA; 179649 MW; C06DB176AC9C8454 CRC64;

Query Match 39.5%; Score 1590.5; DB 5; Length 1584;

Best Local Similarity 44.7%; Pred. No. 4.7e-86;

Matches 356; Conservative 139; Mismatches 213; Indels 89; Gaps 19;

QY 6 NIKVVVRFPNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAPA 65
DB 3 SYKVAVRVPFNQREISNTSKCVLQVNGNTTI-----NGHSINKNENSFNFD 50
QY 66 RSYWSEFNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQSGSKSYMMGYG-- 123
DB 51 HSYWSEFNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQSGSKSYMMGYG-- 110
QY 124 -EHGVIPRICQDMFRINELQDKNITCTVEVSYLEIYNERNVRLDLPSTKGNLKYREHP 182
DB 111 DEMGIIPRLCNDLFAIDN-NNDKDVQYSVEYSVEIYCEYKVLNPNSSGNLVRREHP 169
QY 183 STGPYVEDLAKLVRSFOELENLMDGKNKARTVAATNMNETSSRSHAVFTLTITQKWHDE 242
DB 170 LLGPYVDDLTQKAVCSYHDICNLMDEGNKARTVAATNMNETSSRSHAVFTIILTKRHCA 229
QY 243 ETKMDTEKAKISLVDLGASERATSGATCARLKEGAENRSLTGLRVIAALADMS 302
DB 230 DSNLDTKHSKISLVDLGASERANSTGAGQRLKAGANINKSLTTLGLVISKLAESTKK 289
QY 303 OKKNO-LVYPRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTRYADSAKRKN 361
DB 290 KKSNGKVIYRDSVLTWLLRNLGNSKTAMLAALSPADINFDLSTLRYADRAKIVC 349
QY 362 HAVYNEDPNARMIRELEKEELAQLRKSKLOSSGGGGGAGSGGPGVEESYPPDTPLEKQIVS 421
DB 350 QAVVNEDPNALIRELNEEVKLRHILKDKG-----IDVTD 385
QY 422 IQOPDATYKKMSKA-----EIVEQLNQSEKLYRDLNOTWEEKLAKTEITHKEAALDEL 476
DB 422 IQOPDATYKKMSKA-----EIVEQLNQSEKLYRDLNOTWEEKLAKTEITHKEAALDEL 476

Db	112	EESQAGIIPTOC	EELFKIND-NCNEDMSVSVS-SYMEIYCERVRLDLPNKNKGNLRVRE	169
QY	181	HPSTGPGV	EDLAKLVRSFOETENIMDEGNKARTVAATNMNNETSSRHAVFTLTUTQKH	240
Db	170	HPLGLPGV	EDLSKLVTSYTDADLMADGNKARTVAATNMNNETSSRHAVFTIVTQKK	229
QY	241	DEETKMDTKVAKIS	LDVLAGEERATSTGATGARKLKEGAENRSLSITGLTRVTAALADKSS	300
Db	230	DPENLST	ETKVTKISLDVLAGEERADSTGAKGTRLKEGNNLNSLTITLKGVISALAEVSK	289
QY	301	GQKKKQLVPYR	SDVFWLLKDSLGNSMTAAIISPADINPEETL--STLRYADSAGR	358
Db	290	-KKKKTDFIPYR	SDVFWLLKDSLGNSMTAAIISPADINPEETL--STLRYADSAGR	348
QY	359	IKNHAVVND	PNAMRIELKEELAOLRSLQSSGGG-----	394
Db	349	IKCNVAVIN	EDPNAKLVELKEEVTRLDLRAOGLGDIIDNLKDFQNNKHYLLASENQR	408
QY	395	-GGAGSGG	GPVSESPDPTPLEKQ-----IVSIQOPDATVKKMSKAETVQOLNQSEKLY	448
Db	409	PGNFSTAS	MGSLTSS-FSSCSLSNQAGLTSVSIQ--ERIMSTPGGEEAIERLKESEKII	465
QY	449	RDLNQTWEEK	IAKTEETHKREAALEELGISIEK--GFGPGYHSKEMPHLVNLSDDPLA	506
Db	466	AELNTEWEEK	LRKTEAIRMEREALLAEMGVAIREDGGTLGVFSPKTPQVWNLNEDPLMS	525
QY	507	ECLVYNIK	QGTR-VGNWQDTQAIRLNSKILK-EHCETENVDN-----VTVIVNEK	559
Db	526	ECLLYIK	GTGFGQADARRQDIVLSGAHTKEEHCFISERNTGTEVITLPEPCR	585
QY	560	AAVMVNGVR	IDKPRLSRGYRIILGDPHIFRPNHPEARAEQOSSLRHSVTNSQLGSP	619
Db	586	SETVYNGK	RVAPQLSSGTVSSWKNHVRPNHPQARAEREK-----	629
QY	620	APGHRD	TLKAGSDAGDSRSDSLPHFRGKDSDFWARREASAILGLQOK	672
Db	630	-----	-----TPSAETPSEVDMTAQRELLEK-OGIDMK	658
RESULT	10			
Q9NG02				
IC	Q9NG02	PRELIMINARY;	PRT: 2205 AA.	
AD	Q9NG02:			
DT	01-OCT-2000	(TRENBLurel. 15, Created)		
DT	01-OCT-2000	(TRENBLurel. 15, Last sequence update)		
DT	01-OCT-2000	(TRENBLurel. 15, Last annotation update)		
DE	KINESIN	UNC104/KIF1A HOMOLOG.		
GN	UNC104.			
OS	Dictyostellium	discoideum (Slime mold).		
OC	Eukaryota;	Dictyostellia;	Dictyostellium.	
OX	NCBI_TaxID=44689;			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=20014990;	PubMed=10545495;		
RX	Pollock N., de Hostos E.L., Turk C.W., Vale R.D.;			
RT	"Reconstitution of membrane transport powered by a novel dimeric			
RT	kinesin motor of the Unc104/KIF1A family purified from			
RT	Dictyostellium."			
RL	J. Cell Biol. 147:493-506(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Pollock N., Vale R.D.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF245277; AAF63384.1;			
SQ	SEQUENCE 2205 AA; 248001 MW; 02C5101E9D61C9ED CRC64;			
Query Match	37.9%;	Score 1525.5;	DB 5; Length 2205;	
Best Local Similarity	45.1%;	Pred. No. 5.6e-82;		
Matches 333;	Conservative 118;	Mismatches 167;	Indels 121;	Gaps 13;
QY	6	NIXWVVRV	RPENRGAACIVRMESGNTITLTPPGAAEAKRSGKGTIMDGKAFAD	65
		:-:	:-:	:-:

```
Db 2 NVQAVRVRFNSREKERNAEILVQMNNKSTILTRPSAL--RANPLAAPTADDEKSESF 59
Qy 66 RSYRSFDKAPNARYAQEDLFDGLVPLLDNAFKYNNCIFAYGQTGSGKSYSMGYGREH 125
Db 60 YSYWSDNSDFHAFSQSVTYNDLGEVLKNAWDGFCNFCSIFAYGQTGSGKSYSMGYGEEK 119
Qy 126 GVIPRICODMFRINELQKDN--LTCVVEVSYLBIYNERVDRLLNPSTK--GNLKVREH 181
Db 120 GIPLICEELFORQSTSPNSNEQTIYTVSYMIYNEKVKDLNNNNKTGGUKVRNN 179
Qy 182 PSTGPVEDLAKLVRSFQEIENLMDGKNKARTVAATNNNETSSRSHAVFTLTITQKWH 241
Db 180 PSTGPVEDLSKLVKSFSEIDMLMDGSKARTVASTNNATSSRSHAVFTIVFTQSKID 239
Qy 242 EETKMDTEKAKISLVLDLAGSERATSTGATGARLKEGAEINRSLSLGRVIAALADMSSG 301
Db 240 KTRCTAIDRVSKISLVLDLAGSERANSTGATGVRLEKAGANINKSLSTGLKVISALAE 298
Qy 302 KOKKNQLYPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSARKIKN 361
Db 299 --SKAVFPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSARKIKT 357
Qy 362 HAVVNEDNARMIRELKEALQRLSKLQSSGGGGGSGGSGGPPVEEYPPDTPLEKQIVS 421
Db 358 VAVVNEDAQSKLIRELQGEVERLRAMDOGG-----QYHANDSKLMN----- 399
Qy 422 IQOPDATVKKMSKAEIVQLNOSEKLYRDLOTWEEKLAKTEEIHKEREAALEELGISIE 481
Db 400 --SDVDETYSLTN-----BKIEQYELKMAELNKSWEELSEAEAIREDRMAALKDMGVAIK 453
Qy 482 KGFVGPYHSKEMPHLVNLSDDPLLAECVLYNKKPQTRVGNVQDTQAEIRLNGSKILKE 541
Db 454 V-----VSSPHILNIDPLMSESIYVYKGEKTRIGRSDSEIPDIIILNGIHNHKE 506
Qy 542 HCTFENVNVTVIPN-----TKSPVDQI-----MDYDFALNELAS--IQGLTAMSKH 655
Db 507 HCIPENINGKVLISFNNFNNNNKENSSTPTTSSKSPKPKSEKKEKNNDDDDGGEK 566
Qy 558 --EKAAMVNGVRIDKPRLSRGYRIILGDHFIRFNHPEARAEQOSLLRHSVTNSQ 615
Db 567 KLDRSYIYVNGVEINPKPTILTGNRVILGNHIFRNFNPEEAIAKAREGNTGTTGIVSS- 625
Qy 616 LGSAPAGRHRTLSKAGSDADGSDSPHFRGKSDWFAREAAASAILGLDQKISH 675
Db 626 -----TKSPVDQI-----MDYDFALNELAS--IQGLTAMSKH 655
Qy 676 LTD-----DELDAFDVQ 689
Db 656 INDKOYKKQMAUYDQIR 674

RESULT 11
ID O9NOT8 PRELIMINARY; PRT; 1826 AA.
AC O9NOT8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE KINESIN-LIKE PROTEIN GAKIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hanada T., Lin L., Tibaldi E.V., Reinherz E.L., Chishti A.H.;
RT "GAKIN: a novel kinesin-like protein associates with the human homolog
of the Drosophila discs large tumor suppressor in T lymphocytes.";
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AF279865; AAF81263.1;
SQ SEQUENCE 1826 AA; 202665 MW; C614E7F3A89E89ED CRC64;
```

```
Query Match 35.0%; Score 1412; DB 4; Length 1826;
Best Local Similarity 44.7%; Pred. No. 2.3e-75;
Matches 318; Conservative 114; Mismatches 193; Indels 86; Gaps 17;

Qy 3 GGGNIKVVVRVPENAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIIMDGPKAF 62
Db 2 GDSKVAVAVIRPNNRRETDLHTKCVVDVANKVILNPVNTNLSKGDARGQ-----PKCF 56
Qy 63 APDRSYWFDKNA--PNYARQEDLFDGLVPLLDNAFKYNNCIFAYGQTGSGKSYSMGY 121
Db 57 AYDHCFTWSDSESVKEKYGQDIFVKCLGENILQNAFDGYNACIFAYGQTGSGKSYTMGT 116
Qy 122 GKEGVIPRICODMFRINELQKDN--LTCVVEVSYLBIYNERVDRLLNP--STGKNLKV 178
Db 117 ADQPLGILPRCSGLFER---TQKEENEQSFKEVSYMEIYNEKVRDLDLPKGSQRLTKV 173
Qy 179 REHPSTGTYVEDLAKLVRSFQEIENLMDGKNKARTVAATNNNETSSRSHAVFTLTITQK 238
Db 174 REHSLVGPYVDGLSKLAATSYKDIESLMDGSKARTVASTNNATSSRSHAVLKTILTH 233
Qy 239 WHDETKMDTEKAKISLVLDLAGSERATSTGATGARLKEGAEINRSLSLGRVIAALADM 298
Db 234 LYDAKSGTSGEKVGLSLVDLAGSERATKTGAAGDRLKEGINSINSLTTLGLVISALADQ 293
Qy 299 SSGKOKKNQLYPYRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSARK 358
Db 294 SAGK--NKNKFPYRDSVLTLLKDSLGNSMTAMVATVSPADNYDETLSLRYADRAKH 352
Qy 359 IKNHAVVNEDNARMIRELKEALQRLSKLQSSGGGGGSGGSGGPPVEEYPPDTPLEKQ 418
Db 353 IVNNAVVNEDNARMIRDLREEVEKLRLQLYKA----- 385
Qy 419 IVSIOQDATVKKMSKAEIVQLNOSEKLYRDLOTWEEKLAKTEEIHKEREAALEELGI 478
Db 386 -----EAMKSPELKORLESEKLIQEMTWEKLRKTEIAQEROKQLESIGI 434
Qy 479 SIEKGFVGPYHSKEMPHLVNLSDDPLLAECVLYNKKPQTRVGNVQDTQAEIRLNGSKI 538
Db 435 SLOSS--GIKVDGDDCKFLVNLNADPALNELLYVYLKE--HTLIGSANSQ---DIQLCGMI 488
Qy 539 LKEHCTFE--NVDNVTVIIPNEKAAMVNGVRIDKPRLSRGYRIILGDHFIRFNHPE--EE 596
Db 489 LPEHCLIIDITSEGQVMLTPQKNTRTFVNGSSVSSPIQLHHGDRILWGNHFRNLNPKK 548
Qy 597 ARAERQEGS--LLRHSVTNSQLGSPAPGRHRTLSKAGSDADGSDSPD--SPLPHFRGKD 652
Db 549 KKAEREDEQDPSMKNNSSQOL-----DVGDSSESSEVSNF----- 587
Qy 653 SDWFAREAAASAILGLDQKISHLTDDELDAFDVQKARAVRGLVEDNE 703
Db 588 --NYEAOEMVTMKALGSNDPMQSI--NSLEOQHEBEKRSALERQRLMYEHE 636

RESULT 12
O01349 PRELIMINARY; PRT; 1921 AA.
ID O01349
AC O01349;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE KINESIN-73.
DE KINESIN-73.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97188425; PubMed=9037010;
RA Li H.P., Liu Z.M., Nirenberg M.;
RT "Kinesin-73 in the nervous system of Drosophila embryos.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:1086-1091(1997).
DR EMBL; U81788; AAB50404.1; -;
```



```
Db 59 CFYSLNPEDENAFQETVDCVGRILDNAFGYNACIFAYGTCGSKSYTMGTQESKG 118
Qy 127 VIPRICQDMFRINELQKDKNLTCTVEYSYLEIYNERVDDLNPSTGKNTLKVREHPSTG 185
Db 119 IIPRLCDQLFSAIAN-KSTPELMYKVEYSYMEIYNEKVDHLLDPKPKQSLKVRHHVMG 177
Qy 186 PYVEDLAKLVRSFOEINLMDGKNKARTVAATNNETSSRSRHAVFTLTQKHDEETK 245
Db 178 PYVDGLSLQVTSYODIDNLTGKNSRTVAATNNNAESSRHAVFVVLTAQILTDQATG 237
Qy 246 MDTEKVAKISLVDLAGSRATSTGATGARLKEGAINSLTGLRVIAALADMSSGCKK 305
Db 238 VSGEKVSRMSLVDLAGSRVKTGAVGRLKEGNSINIKSLTGLVLSKLDQSNKGKSG 297
Qy 306 N-QLVPPYRDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLRVYADSAKRKNH 364
Db 298 NDKFVPPYRDSVLTWLLKDNLGNSRTVMVATISPSADNYEETLSTLRVADRAKRI 357
Qy 365 VNEDPNARMIRLEKEELAQLSKLQSSGGGGGSGGPGVEESYPPDTPLEKQIVSIQQ 424
Db 358 VNEDPNARMIRLEKEELAQLSKLQSSGGGGGSGGPGVEESYPPDTPLEKQIVSIQQ 389
Qy 425 PDATVKKMSKAEIYEQLNQSEKLYRDLNQTWEELAKTEETHKEREAALEELGSIKGF 484
Db 390 -----GDVQDKLAESNLKMQISQWEEKLYKTERIQNERQQALEKGLSVQAS- 438
Qy 485 VGPYHSEKMPHLVNLSDPDLAELVYNIKPGQTRVGNVNDTQAEIRLNGSKILKEHCT 544
Db 439 -GIKVEKNKYLVNLNADSNELVLYLKDTLIGRTISQQPDQLGSLGIGQIEHC 497
Qy 545 FENVNVTYIYVNEKAAVYNGVRIDKTRLSGVRILGDFHIFRPNHPE----- 595
Db 498 ITIEDSGLYMEFVGARCFVNGSAAVETPLQNGDRILLGNHHFRVNSPKNSNTSMCAS 557
Qy 596 --EAAERQEQSLRLSVNTSOLGS 618
Db 558 EPQTPAQILIDYINFADEIMQNELSN 582

RESULT 14
Q20888 ID Q20888 PRELIMINARY; PRG; 1576 AA.
AC Q20888
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE SIMILAR TO KINESIN-RELATED PROTEIN. NCBI GI: 1109842.
GN F56E3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditioidae;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Spratt J., Wohlman P., III of C.
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.;
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
```

```
RA Gattung S., Wu X.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41536; AAB52613.1; -.
DR HSSP; P33176; 18G2
DR INTERPRO; IPR001752; -.
DR PFAM; PF00225; Kinesin; 2.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1576 AA; 177409 MW; 385D820EDF588B66 CRC64;

Query Match 31.9%; Score 1284.5; DB 5; Length 1576;
Best Local Similarity 38.2%; Pred. No. 7.1e-68;
Matches 323; Conservative 120; Mismatches 217; Indels 185; Gaps 23;

Qy 7 IKVVYVRRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMGPKAFADR 66
Db 10 VKVAIRVRPFNKRRLDLTKSVVRIQKQCVLHHP--TEEK-----NSKTFTH 57

Qy 67 SYWSPDKNAPYARQEDLFQDLGVPLLDNAFGYNACIFAYGTCGSKSYTMGTQESKG 126
Db 58 SFCSTDPHSYDFASQETVSYHLGSGVWENAFSGYNACIFAYGTCGSKSYTMGTQDQPG 117

Qy 127 VIPRICQDMFRINELQKDKNLTCTVEYSYLEIYNERVDDLNPSTGKNTLKVREHPSTG 185
Db 118 IIPRVNDIFTRIQE-TSNSSLSFKVEYSYMEIYNERVDDLDPKSKKALKVREHKILG 176

Qy 186 PYVEDLAKLVRSFOEINLMDGKNKARTVAATNNETSSRSRHAVFTLTQKHDEETK 245
Db 177 PMVDGLSLAVNSFRQISNLLGEGNKSRTVAATNNNAESSRHAVFSLIVTQTLHDLENG 236

Qy 246 MDTEKVAKISLVDLAGSRATSTGATGARLKEGAIN----- 282
Db 237 FSGEKVAKISLVDLAGSRAGTGAVGRLKEEGGNKLNKLSIFLRNDLEKKIDPKFSAD 296

Qy 283 -----RSLSTLGRVITAAADMSSGCKKQKQNLVPPYRDSVLTWLLKDSLGNSMTA 331
Db 297 VYVFINOKLDFRSLTTLGWNVISALAEARN--KKDKFFPYRDSVLTWLLKDSLGNSRTV 353

Qy 332 MTAISPADINFEETLSTLRVYADSAKRKNHVNEDPNARMIRLEKELQALRSKLQSS 391
Db 354 MTATLSPAADNVEETLSTLRVADRAKRIIVHAIINEDPNARVIRELREBEVETLR----- 407

Qy 392 GGGGGGAGSGGPGVEESYPPDTPLEKQIVSIQQPATVKKMSKAEIYVQLNGSEKLYRDL 451
Db 408 -----MQITQTKKEAETE-----ELRERLAESEKRLVQAM 437

Qy 452 NOTWEEKLAKTEETHKEREAALEELGISTEGFVGYPKSKEMPHLVNLSDPDLAELVY 511
Db 438 NKSWEERLKEHCTDLNKRKQKDLTEIGISIESS--GIKVEKDRFYLVNNADSLNELLYV 495

Qy 512 NIKPGQTRVGN-----VNQDT-----OAEIRLNGSKILKEHC--TFEN 547
Db 496 YIN-GSAIGNSELETSRDSGLSMTCSDSSRRDDKERTSVLRGLGIMRRHAKMTVEE 554

Qy 548 VDNVVT--IYV-NEKAAVYNGVRIDKTRLSGVRILGDFHIFRPNHPEARERQEQ 604
```



```
Db 555 YGRLRLFPAPMSSECRICVNGKQITERLLRNGNRLLVGMNHFVKVNC-----KVMDMEQ 611
Qy 605 SLRLHVSNTNSQLGSPAPGRDRTLSKAGSDADGSDSDSLPHFPGKDSDFWYARREAS 664
Db 612 SIMEDSTFDY-----NDAMHEVDANPI-----SS 637
Qy 665 AILGLDQKISHLTDELDFDVOKARAVRGGLVEDNEDSDS--QSSFPVRYKYSNGT 722
Db 638 AV---DOYMESVT-----LKHQEDKKAALQQYEAFEKYIQSUT 673
Qy 723 IDNFSLDTAITMPG-----TPRSDDDGDALFFGDKKSKQDASNVDEELRQQQAQMEELK 778
Db 674 AGGFTSPMT-PGFCPLPTITPGLPPFPFPPANPQSVKSKFFYNAQKEMPASLK 732
Qy 779 TAKQE 783
Db 733 RLKAD 737

RESULT 15
Q15058 PRELIMINARY; PRT: 1648 AA.
AC Q15058;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE KIAA0042 PROTEIN.
GN KIAA0042.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBL_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051396; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
DR EMBL; D26361; BAA05392.1; -.
DR HSP; P17119; 3KAR.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR PFAM; PF00225; kinesin: 1.
DR PFAM; PF00498; FHA: 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00467; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1648 AA; 186490 MW; FBI423668A7B79D7 CRC64;
```

```
Query Match 29.1%; Score 1174.5; DB 4; Length 1648;
Best Local Similarity 34.8%; Pred. No. 2.6e-61;
Matches 281; Conservative 152; Mismatches 239; Indels 135; Gaps 17;

Qy 7 IKVVRVRFNAREIDRGAICVIRMEGNOTILTPPPGAEEKARKSGKTIIMGPKAFADR 66
Db 359 VTAVVRVPFTKREKIEKASQVVMFSGKIEITVEHP-----DTQVYN---FIYDV 405

Qy 67 SYWSPDKNAPNARQEDLFDGLVPLLDNAPKYNVICFAYGOTGSGKSYSMGMGKKEHG 126
Db 406 SFWSPDECHPYASQTTVYEKLAAPLLERAFEGENTCLFAYCQTGSGKSYTMGFSPEPG 465

Qy 127 VIPRCQMFRRINEKQKNKNTCTVSVSYLEYINVERVDLL-----NPSTGNLKVREH 181
Db 466 IIPRCEDLFSQVARKO-QTEYSYHIEMSFYVNEKIHDLLVCKDENGORKQPLVREH 524

Qy 182 PSTGYPVEDLAKLVRSFOEINLMDENKARTVAATNMNETSSRSHAVFTLTQKWHID 241
```

Search completed: April 25, 2001, 10:18:02
Job time: 308 sec

```
Db 525 PVYGPYVEALSMNIVSSVADIQSMLELGNKQRATAATGMNDKSSRSHSVFTLVMTQTKTE 584
Qy 242 --EETKMDTEKVAKISLDVLAGSERATSTGATCARLKEGAENESLSTLGRVIAALADMS 299
Db 585 FYEGEEHDHRTSRINLIDLAGSECSCTAHTNDRLKEGVSINKSLITLCKVISALSQA 644
Qy 300 SKCKNQOLVPYRDSVTLWLKSLGNSMTAMIAISPADINFEETLSTLRYSADAKRI 359
Db 645 N--QRSVFIPYRESVTLWLKSLGNSKTMATISPAASNTEETLSTLRYSANQARLI 701
Qy 360 KNHAVYVEDNAPMIRELKEALQRLSKLQSGGGGGGAGGSGGPVEESYPPDTPLEKQI 419
Db 702 VNIKVNEDNNAKLRELKAEIAKLKAQRNS----- 733
Qy 420 VSIQOPATVYKMSKAEIVE--OLNOSKLYRDLNQTWEEKLAKTEEIHKREAALEEL 476
Db 734 ---RNIDPERYRLCROEITSLRMKLHQOERDMAEMQRYMKEKPEQAEKRKLQETKELQA 790
Qy 477 GISIEKGFVGPYHSEMPHLYNLSDPLLAECILVYNIKPGOTRVGNVNOTQAEIRLNGS 536
Db 791 GIMFQM-----DNHLPNLVNLNEDPQLSEMLLYWIKEGTTVGYKYPNSSHDIQLSGV 843
Qy 537 KILKECHTFENVNVVTVPNKAAVWVNGVRIDKPTRLRSGYRIILGDFHIFRPNHPEE 596
Db 844 LIADDHCTIKNFGGTVPYGEAKTYVNGKHILEITVLRHGDRVLGGDHYFRFNHPVE 903
Qy 597 ARAERKQQLRLRHVSNTNSQLGSPAPGRDRTLSKAGSDADGSDSDSLPHFPGKDSDFW 656
Db 904 VQKGR-----PSGRDPTISEGPKDFE-FANELLMAQRSOLEAEIK 944
Qy 657 YARPEASAILGLDQKISHLTDELD---ALFDDVQKARAVRRLGVLVEDNEDSDSSPPV 713
Db 945 EAOLKAKEEMHOGIQTAKEMAQQLSSQKAAYE--SKIKALEAELREESQKMKQEI--- 999
Qy 714 RDKYMSNGTIDNFSLDTAITMPGTTPRSDDDGDALFFGDKKSKQ-----DASNVVDVEE 765
Db 1000 -NNQKANHKEEL-----EKAKQHLQEQEIVVNNKRLKLEMET 1033
Qy 766 LROQQ-----AQMEALKTAQOE 783
Db 1034 LATQOALEDSIRHARILEALETEBKQ 1060
```




GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:12:41 ; Search time 56.07 Seconds
(without alignments)
363.961 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357
Perfect score: 1834

Sequence: 1 MSGGNIKVVVRPFNARE.....PADINFEETLSRLYADSAK 357

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1834	100.0	784	20 Y06618	Thermomyces lanugin
2	1119	61.0	1816	21 B36227	Human kinesin-like
3	1117	60.9	1103	21 Y51328	Human KLMP protei
4	913.5	49.8	503	21 B63190	Human secreted pro
5	911.5	49.7	504	21 B63189	Gene 5 human secre
6	705.5	38.5	955	15 R57365	K39 polypeptide of
7	705.5	38.5	955	17 W03691	Leishmania chagasi
8	615.5	33.6	2954	20 W01632	Amino acid sequenc
9	582	31.7	411	19 W72745	Drosophila kinesin
10	582	31.7	441	19 W72744	Drosophila kinesin
11	582	31.7	975	19 W72746	Drosophila kinesin

12	542	29.6	1269	21 Y77955	A. thaliana enviro
13	515	28.1	1518	21 G40075	Arabidopsis thalia
14	515	28.1	1662	21 G31282	Arabidopsis thalia
15	511.5	27.9	1460	21 G40077	Arabidopsis thalia
16	511.5	27.9	1462	21 G40076	Arabidopsis thalia
17	511.5	27.9	1604	21 G31284	Arabidopsis thalia
18	511.5	27.9	1606	21 G31283	Arabidopsis thalia
19	510	27.8	469	21 G21665	Arabidopsis thalia
20	495.5	27.0	452	21 G21666	Arabidopsis thalia
21	472	25.7	398	21 G21667	Arabidopsis thalia
22	464.5	25.3	324	19 W70235	Leishmania antigen
23	460	25.1	460	21 B56650	Human prostate can
24	433.5	23.6	679	20 W88456	Human kinesin-rela
25	433.5	23.6	730	21 Y49949	Xenopus laevis kin
26	433.5	23.6	1034	21 G31112	Arabidopsis thalia
27	433.5	23.6	1069	21 G31111	Arabidopsis thalia
28	433.5	23.6	1121	21 G31110	Arabidopsis thalia
29	424.5	23.1	829	21 G31117	Arabidopsis thalia
30	424.5	23.1	834	21 G31116	Arabidopsis thalia
31	404	22.0	790	21 G41923	Arabidopsis thalia
32	404	22.0	794	21 G41922	Arabidopsis thalia
33	404	22.0	814	21 G41921	Arabidopsis thalia
34	403	22.0	147	21 G02949	Human secreted pro
35	379.5	20.7	410	21 B56496	Human prostate can
36	324	17.7	959	21 Y77944	A. thaliana enviro
37	317	17.3	726	21 G31118	Arabidopsis thalia
38	302	16.5	154	21 B40661	Human ORFX ORF425
39	260	14.2	1201	20 W90345	Drosophila sp. Cos
40	211.5	11.5	243	21 Y92345	Human cancer assoc
41	163	8.9	92	21 B41721	Human ORFX ORF1485
42	146	8.0	482	21 B56609	Human prostate can
43	107	5.8	338	21 G41653	Arabidopsis thalia
44	107	5.8	338	21 G49380	Arabidopsis thalia
45	107	5.8	386	21 G49379	Arabidopsis thalia

ALIGNMENTS

RESULT 1
Y06618
ID Y06618 standard; Protein; 784 AA.
XX AC Y06618:
XX DT 26-OCT-1999 (first entry)
XX DE Thermomyces lanuginosus kinesin motor protein TL-gamma.
XX TL-gamma; kinesin; motor protein; microtubule; unc-104; Infection;
XX KW neurodegenerative disease; Alzheimer's disease;
XX KW Parkinson's disease; Huntington's disease;
XX KW amyotrophic lateral sclerosis.
XX OS Thermomyces lanuginosus.
XX PN WO9937659-A1.
XX PD 29-JUL-1999.
XX PF 22-JAN-1999; 99WO-US01355.
XX PR 23-JAN-1998; 98US-0072361.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Goldstein LSB, Sakowicz R;
XX WPI: 1999-493950/41.
XX N-PSDB: X87656.
XX PT New nucleic acid encoding microtubule motor protein, used for diagnosis of fungal infection and neurodegenerative disease


```
Best Local Similarity 61.9%; Pred. No. 8.6e-82;
Matches 190; Conservative 40; Mismatches 64; Indels 13; Gaps 6;

QY 58 GP-KAFAPDRSWSFDKNA-PHYAQEDLFDQLGVPLLDNAFGYNNCIFAYGQTSGSKS 115
Db 9 gpaevfaydhtcwsmdesvkekyagdvfkclgenilqnaixgnacifaygqtsgks 68

QY 116 YSMWVGKEHGVIPRICQDMFRRI----NELQDKNLCTVEVSYLEIYNERVRLDLP- 170
Db 69 ymmgtadpgliplrclsglfertkqegneqsfk-----vevsymelynekrvdlldpk 123

QY 171 STKGNLKVREHSTGPYYVEDLAKLVYRSFQETENLMDENKARTVAATNMNFTSSSHAV 230
Db 124 gsrqtlkvrehavlpvydvglsklavtsykdlieslmsegknkrtvaatnmneessrshav 183

QY 231 FTLTLTQKWHDETQNDKYEKAKISLVLDLAGSERATSGCATGARLKEGAEINRSLSLIGR 290
Db 184 fktitlhtldyvksgtsgkvgkxslvdlxgseratktagaagdrllkegsnlnksittlgl 243

QY 291 VTAALADMSSGKOKKKNOLVPYRDSVLTWLLKDSLGNSTMTAAISPADINFEETLSLT 350
Db 244 visaladgaqk-sxnkfpyrdsvtlwlkdslgnskxamvatvspaadhydecisc 302

QY 351 RYADSAK 357
Db 303 ryadrak 309

RESULT 5
B63189
ID B63189 standard; Protein; 504 AA.
AC B63189;
XX
DT 26-MAR-2001 (first entry)
DE
DE 5 human secreted protein homologous amino acid sequence #115.
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
XX
OS Drosophila melanogaster.
PN
PN WO200061629-A1.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US09071.
XX
XX 09-APR-1999; 90US-0128694.
XX
XX 20-JAN-2000; 2000US-0176931.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-647420/62.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -
XX
XX Disclosure; Page 487-488; 533pp; English.
XX
XX F22373 to F22421 encode the human secreted proteins given in B63134 to
XX B63182. B63183 to B63231 represent more human secreted proteins and
XX
```

```
polypeptides homologous to them. Human secreted proteins have activities
based on the tissues and cells the genes are expressed in. Examples of
activities include: immunosuppressive; antiarthritic; antirheumatic;
antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
neotropic; neuroprotective; antibacterial; virucide; fungicide; and
ophthalmological. The polynucleotides and proteins can be used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also
used in diagnosing a pathological condition or susceptibility to a
pathological condition. Disorders which are diagnosed or treated include
autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
angiogenesis, nervous system disorders e.g. Alzheimer's disease,
infections caused by bacteria, viruses and fungi and ocular disorders
e.g. corneal infection. The polypeptides can also be used to aid wound
healing and epithelial cell proliferation, to prevent skin aging due to
sunburn, to maintain organs before transplantation, for supporting cell
culture of primary tissues, to regenerate tissues and in chemotaxis. The
polypeptides can also be used as a food additive or preservative to
increase or decrease storage capabilities. F22364 to F22372 and B63133
represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 504 AA;

Query Match 49.7%; Score 911.5; DB 21; Length 504;
Best Local Similarity 61.5%; Pred. No. 1.4e-81;
Matches 185; Conservative 43; Mismatches 70; Indels 3; Gaps 3;

QY 59 PKAFAPDRSWSFDKNA-PHYAQEDLFDQLGVPLLDNAFGYNNCIFAYGQTSGSKSYSM 118
Db 8 pktfahdchfslnpedenfasqetvdcvgrgildnaifgynacifaygqtsgsksygm 67

QY 119 MGYGKEHGVIPRICQDMFRRIINELQDKNLCTVEVSYLEIYNERVRLDLPNS-TKGNLK 177
Db 68 mgtqeskgliplrclqglfsaian-kstpelmykvevsymelynekvhdlldpkpnkqslk 126

QY 178 VREHSTGPYYVEDLAKLVYRSFQETENLMDENKARTVAATNMNFTSSSHAVFTLTQ 237
Db 127 vrehvmvgpydvglsglavtsygdldnImtegnkrtvaatnmnaessrshavfsvltq 186

QY 238 KWHDEETKMDTEKVKAKISLVLDLAGSERATSGCATGARLKEGAEINRSLSLIGR 297
Db 187 iltqatgvsgekvmslvdlagseravktgavdrlkegsnlnksltlglvsklad 246

QY 298 MSSGKOKKN-QLVPYRDSVLTWLLKDSLGNSTMTAAISPADINFEETLSLTRYADSA 356
Db 247 qsgnkksgndkfpyrdsvtlwlkdslnlgnsrtvmvatvatspsadnyeetlsltryadra 306

QY 357 K 357
Db 307 k 307

RESULT 6
R57365
ID R57365 standard; Protein; 955 AA.
XX
XX R57365;
XX
XX AC
XX
XX 08-MAR-1995 (first entry)
XX
XX K39 polypeptide of Leishmania chagasi.
XX
XX Parasite; leishmaniasis; infection; repeat unit; vaccine; antigen;
XX Leishmania chagasi.
XX
XX OS Leishmania chagasi.
XX
XX PN WO9416331-A.
XX
XX 21-JUL-1994.
XX
```


SQ	Sequence	411 AA;	
	Query Match	31.7%; Score 582; DB 19; Length 411;	
	Best Local Similarity	40.0%; Pred. No. 4.6e-49;	
	Matches 142; Conservative	59; Mismatches 116; Indels 38; Gaps 10;	
QY	6 NIKVVVRPFNAREIDRGAKCIVRMEGNQTLTPPGAEERKARKSGKTMGPKAFAPD	65	
DB	12 sikvccfrlpndseekagkfvvkf-----pnvveenc-----islagkvyllfd	56	
QY	66 RSWSFDPKNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQTGSGKYSVM---MGYG	122	
DB	57 kvf-----kpn-asqekyvneaaaksvtdvlagyngtifaygqtsqgkthtmegvlgds	109	
QY	123 KEHGVTPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDNLNPKGNLKVREHP	182	
DB	110 vkqgiiprivndifnhiyam--evnlefthikvsvyelymdkirdldvs-kvnlsvhecd	166	
QY	183 STGPVYEDLAKLVRSFQEIENLMDGNKARTVAATNMNETSSRSRAVFTLTQKHDE	242	
DB	167 nrpvykgaterfvsspedvfeveegksnrhnavtmnehssrshsvflnvkqenlen	226	
QY	243 ETKMDTEKVAKISLVDLAGSRATSTGATGARLKEGAETNRSLSLTGRVIAALADMSSGK	302	
DB	227 qkklis----gklylvdlagsekvsktgagtvdeaknkinkalsalgnvisalad---gn	279	
QY	303 QKNQOLVPRYDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLYADSAK	357	
DB	280 kth---ipyrdskltrilqeslggnartttivccspasfnesetkstdldfgrak	331	
RESULT	10		
ID	W2744		
XX	W2744 standard; protein; 441 AA.		
AC	W2744;		
XX	11-JAN-1999 (first entry)		
DE	Drosophila kinesin N-terminal 441 amino acid residues.		
XX			
KW	Drosophila; kinesin; separation; hybridisation; target site;		
KW	complex mixture; motor protein; actively transported; separated;		
XX	microtubule.		
OS	Drosophila sp.		
XX			
PN	US5830659-A.		
PD	03-NOV-1998.		
XX			
PF	13-SEP-1996; 96US-0713815.		
XX			
PR	13-SEP-1996; 96US-0713815.		
XX			
PA	(UTAH) UNIV UTAH RES FOUND.		
XX			
PI	Stewart RJ;		
XX			
DR	WPI; 1998-609236/51.		
XX			
PT	Separation of selected molecules, e.g. DNA, from complex mixtures -		
PT	uses specific apparatus to allow the selected molecule to bind to		
PT	motor proteins, and be actively transported and separated away along		
PT	microtubules		
XX			
PS	Claim 3; Column 23-26; 24pp; English.		
XX			
CC	A method has been developed of separating a selected molecule from a		
CC	mixture of molecules. The method comprises: (a) a separation device		
CC	comprising a loading reservoir and a receiving reservoir coupled by a		
CC	channel with microtubules immobilised on its surface and aligned parallel		

CC	to a longitudinal axis of the channel; (b) loading the loading reservoir		
CC	with an aqueous solution of the mixture of molecules; (c) adding a motor-		
CC	ligand complex and ATP to the solution, where the motor-ligand		
CC	comprises: (i) a processive motor capable of attaching to the immobilised		
CC	microtubules, and moving in the presence of ATP as source of chemical		
CC	energy, and (ii) a ligand coupled to the motor protein, where the ligand		
CC	is capable of binding the selected molecule, so that the ligand binds the		
CC	selected molecule and the motor protein attaches to the immobilised		
CC	microtubules and transports the bound selected molecules along the		
CC	receiving reservoir; and (d) removing the selected molecule from the		
CC	receiving reservoir. The method and the system are used for the		
CC	separation of specific molecules from complex mixtures. The molecule to		
CC	be separated (e.g. DNA) binds to the motor protein due to the presence of		
CC	the specific binding ligands. Activation of these enables them to travel		
CC	down a preformed channel in a specially made piece of apparatus. They can		
CC	then be removed easily without contaminants of other mixture particles.		
CC	The present sequence represents the N-terminal 441 amino acid residues of		
CC	Drosophila kinesin for use in the method of the invention.		
XX			
SQ	Sequence	441 AA;	
	Query Match	31.7%; Score 582; DB 19; Length 441;	
	Best Local Similarity	40.0%; Pred. No. 5.2e-49;	
	Matches 142; Conservative	59; Mismatches 116; Indels 38; Gaps 10;	
QY	6 NIKVVVRPFNAREIDRGAKCIVRMEGNQTLTPPGAEERKARKSGKTMGPKAFAPD	65	
DB	12 sikvccfrlpndseekagkfvvkf-----pnvveenc-----islagkvyllfd	56	
QY	66 RSWSFDPKNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQTGSGKYSVM---MGYG	122	
DB	57 kvf-----kpn-asqekyvneaaaksvtdvlagyngtifaygqtsqgkthtmegvlgds	109	
QY	123 KEHGVTPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVRLDNLNPKGNLKVREHP	182	
DB	110 vkqgiiprivndifnhiyam--evnlefthikvsvyelymdkirdldvs-kvnlsvhecd	166	
QY	183 STGPVYEDLAKLVRSFQEIENLMDGNKARTVAATNMNETSSRSRAVFTLTQKHDE	242	
DB	167 nrpvykgaterfvsspedvfeveegksnrhnavtmnehssrshsvflnvkqenlen	226	
QY	243 ETKMDTEKVAKISLVDLAGSRATSTGATGARLKEGAETNRSLSLTGRVIAALADMSSGK	302	
DB	227 qkklis----gklylvdlagsekvsktgagtvdeaknkinkalsalgnvisalad---gn	279	
QY	303 QKNQOLVPRYDSVLTWLLKDSLGNSMTAMIAISPADINFEETLSTLYADSAK	357	
DB	280 kth---ipyrdskltrilqeslggnartttivccspasfnesetkstdldfgrak	331	
RESULT	11		
ID	W2746		
XX	W2746 standard; protein; 975 AA.		
AC	W2746;		
XX	11-JAN-1999 (first entry)		
DE	Drosophila kinesin.		
XX			
KW	Drosophila; kinesin; separation; hybridisation; target site;		
KW	complex mixture; motor protein; actively transported; separated;		
XX	microtubule.		
OS	Drosophila sp.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference	557	
XX		/note= "encoded by AAT"	
PN	US5830659-A.		
XX			


```
Db      1106  esidliqcsaargklsfvdliagsservytksgagcqlkeagsinksalsalgdvigal---s 1162
QY      300  SGKQKKQLVPRDVSVLFWLKDSLGNSKMTAMIAISPADINFEETLTSLRYADSAK 357
Db      1163  ag----nghipyrnhkltmlmsdlsignaktimfvnvspaesnldetynsliyasrvr 1216

RESULT  13
G40075  G40075 standard; Protein: 1518 AA.
AC      G40075;
XX      DT 18-OCT-2000 (first entry)
XX      DE Arabidopsis thaliana protein fragment SEQ ID NO: 49676.
XX      KW Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX      OS Arabidopsis thaliana.
XX      PN EP1033405-A2.
XX      PD 06-SEP-2000.
XX      PF 25-FEB-2000; 2000EP-0301439.
XX      PR 25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132484.
PR      06-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      07-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
```

QY	73	KNAPYARQEDLEFDQDLGVPLLDNAFYGNNCIFAYGOTSGKSYSMWG	-----YG 122
Db	91	-----gvkeqmfqlvgaplivencslgfnssvfaygtgsgktytmwgpanglleehlog	144
QY	123	KEHGVTPRICQDMFRINE-----LQDKNLCTVEVSYLEIYNERNVRLDN-	169
Db	145	dqrgltprvferifarikevmvpgvfqvltnkssmlkgn-sitnvghysrfsdwlhy	203
QY	170	-----PSTKGNL-----KVREHPSTGPVVEDLAKUWRSFOE	201
Db	204	islddlqrannrptgpkpkdgkghvhrfslmchqiredvksvgyvenlteeyvknld	263
QY	202	IENLMDGNKARTVAATNMNETSSRSHAVFTLTQKWHDEETKMPTEKVKISLYDLAG	261
Db	264	vsqlikglnrrtgatvntessrshcvftcvvesrcknvaadglssfktsrlnlvdag	323
QY	262	SERATSTGATGARLKKEGAENRSLTLGRVIAALADMS-SGKQKKNQLPYRDSVLTWLL	320
Db	324	serqstgaagerlkkeagnlnrslsqlnlinlaelsqtkprh---ipydrsitfll	380
QY	321	KDSLGNSMTAMIAALSPADINFEETLSTLRYADSAK	357
Db	381	qeslggnaklamvcavspqsrs--efstlrfagrak	415
RESULT 14			
ID	G31282	G31282 standard; Protein; 1662 AA.	
XX	AC	G31282;	
XX	DT	17-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 37541.	
KW	KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	XX	termination sequence.	
OS	XX	Arabidopsis thaliana.	
PN	XX	EP1033405-A2.	
PD	XX	06-SEP-2000.	
PF	XX	25-FEB-2000; 2000EP-0301439.	
XX	PR	25-FEB-1999; 99US-0121825.	
PR	PR	05-MAR-1999; 99US-0123180.	
PR	PR	09-MAR-1999; 99US-0123548.	
PR	PR	23-MAR-1999; 99US-0125788.	
PR	PR	25-MAR-1999; 99US-0126264.	
PR	PR	29-MAR-1999; 99US-0126785.	
PR	PR	01-APR-1999; 99US-0127462.	
PR	PR	06-APR-1999; 99US-0128234.	
PR	PR	08-APR-1999; 99US-0128714.	
PR	PR	16-APR-1999; 99US-0129845.	
PR	PR	19-APR-1999; 99US-0130077.	
PR	PR	21-APR-1999; 99US-0130449.	
PR	PR	23-APR-1999; 99US-0130510.	
PR	PR	28-APR-1999; 99US-0131449.	
PR	PR	30-APR-1999; 99US-0132048.	
PR	PR	30-APR-1999; 99US-0132407.	
PR	PR	04-MAY-1999; 99US-0132484.	
PR	PR	05-MAY-1999; 99US-0132485.	
PR	PR	06-MAY-1999; 99US-0132486.	
PR	PR	06-MAY-1999; 99US-0132487.	
PR	PR	07-MAY-1999; 99US-0132863.	
PR	PR	11-MAY-1999; 99US-0134256.	
PR	PR	14-MAY-1999; 99US-0134218.	
PR	PR	14-MAY-1999; 99US-0134219.	
PR	PR	14-MAY-1999; 99US-0134221.	
Query Match 28.1%; Score 515; DB 21; Length 1518;			
Best Local Similarity 34.5%; Pred. No. 1.7e-41;			
Matches 137; Conservative 63; Mismatches 119; Indels 78; Gaps 12;			
QY	17	NAREIDRGAKC---IVRMGNTILTPPGAEKARKSGKTKMDGPKAFDRSYWSPD	72
Db	41	napiidntstpdhrsmrmk-nplprppnpsnlkrlsaetes-----gfsds-----	90

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142350.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

[illegible]

Search completed: April 25, 2001, 10:12:47
Job time: 343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:13:20 ; Search time 31.19 Seconds
(without alignments)
219.887 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357

Perfect score: 1834

Sequence: 1 MSGGNIKVVVRPFNARE.....PADINFEETLSTLRVADSAK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pap.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pap.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pap.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pap.*
5: /cgn2_6/ptodata/2/1aa/PCIRUS.COMB.pap.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1117	60.9	1103	3	US-09-162-373-1
2	705.5	38.5	955	1	US-08-006-676B-1
3	705.5	38.5	955	1	US-08-282-845-2
4	705.5	38.5	955	2	US-08-428-414A-3
5	705.5	38.5	955	5	PCP-US94-00324-1
6	582	31.7	411	2	US-08-713-815A-4
7	582	31.7	411	2	US-08-713-815A-3
8	100	5.5	545	2	US-08-935-760-4
9	94	5.1	746	4	US-09-149-934-4
10	92	5.0	402	4	US-09-097-767A-26
11	92	5.0	574	3	US-08-732-433-1
12	91	5.0	521	1	US-08-368-281-2
13	91	5.0	581	1	US-08-368-281-4
14	90	4.9	1098	1	US-07-777-715-7
15	90	4.9	1098	1	US-08-170-126-2
16	90	4.9	1098	3	US-08-954-418-2
17	90	4.9	2233	2	US-08-569-853-1
18	90	4.9	2233	2	US-08-569-853-2
19	90	4.9	2233	4	US-08-987-439-1
20	89.5	4.9	544	3	US-08-559-397A-30
21	89.5	4.9	894	3	US-08-362-525-22
22	89.5	4.9	894	4	US-08-971-692-15
23	89.5	4.9	1537	1	US-08-325-267A-2
24	89	4.9	2710	1	US-08-480-604A-6
25	89	4.9	2710	2	US-08-405-496A-6
26	88	4.8	432	2	US-08-705-660-18
27	88	4.8	432	3	US-08-989-045-18

28	86.5	4.7	337	2	US-08-441-857-8	Sequence 8, Appl
29	86.5	4.7	337	4	US-08-193-159-8	Sequence 8, Appl
30	86.5	4.7	387	4	US-09-097-767A-23	Sequence 23, Appl
31	85.5	4.7	862	1	US-08-325-267A-4	Sequence 4, Appl
32	85.5	4.7	1164	2	US-08-589-756-1	Sequence 1, Appl
33	85	4.6	631	1	US-08-487-890A-111	Sequence 111, App
34	85	4.6	631	1	US-08-478-435-111	Sequence 111, App
35	85	4.6	631	2	US-08-337-483-111	Sequence 111, App
36	85	4.6	631	2	US-08-478-373-111	Sequence 111, App
37	85	4.6	631	3	US-08-474-671-111	Sequence 111, App
38	85	4.6	631	3	US-08-483-577A-111	Sequence 111, App
39	85	4.6	936	5	PCT-US93-05944-2	Sequence 2, Appl
40	84.5	4.6	338	4	US-08-441-857-10	Sequence 10, Appl
41	84.5	4.6	338	4	US-08-193-159-10	Sequence 10, Appl
42	84	4.6	1150	2	US-08-589-756-3	Sequence 3, Appl
43	83.5	4.6	565	4	US-08-961-083-218	Sequence 218, App
44	83.5	4.6	954	3	US-09-057-989-2	Sequence 2, Appl
45	83.5	4.6	1657	1	US-08-287-959-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-162-373-1

; Sequence 1, Application US/09162373

; Patent No. 6013454

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Patterson, Chandra

; TITLE OF INVENTION: KINESIN-LIKE MOTOR PROTEIN

; FILE REFERENCE: PP-0593 US

; CURRENT APPLICATION NUMBER: US/09/162,373

; CURRENT FILING DATE: 1998-09-28

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PERL Program

; SEQ ID NO 1

; LENGTH: 1103

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 1281811

US-09-162-373-1

Query Match	60.9%	Score 1117;	DB 3;	Length 1103;
Best Local Similarity	61.6%	Pred. No. 2.7e-106;		
Matches 220;	Conservative 59;	Mismatches 62;	Indels 16;	Gaps 5;
Qy	4	CGNIKVVVRPFNAREIDRGAKCIVRMGNOITLTPPGAEKARKSGKTMDGPKAFA	63	
Db	3	GASVAVRVRRPFNARETSQDAKCVSMQGNITINP-----KQSKDAPKSF	51	
Qy	64	FDRSWSPDKNA-PNYARQEDLFQDLGVLDDNAFGYNNCIFAYGOTGSGKSYMMGYG	122	
Db	52	FDYSWSHTSTEDQFASQOQVRYRDIGEEMLLHAFEGYNVCIFAYGOTGAGKSYTMGRQ	111	
Qy	123	K--EKGIVPRICODMFRRIEQLDKNLTCTVSVSYLYETINRVRLDLPSTKGNKYRE	180	
Db	112	EPGQGVVFPUCEDLFSRVSEHQ-SAQLSYSEVSYMEIYCVRRVRLDLPKPSRGSLRVRE	170	
Qy	181	HPSTGPYVEDLAKLVRSFQETENIMDEGNKARTVAATNMNETSSRSFHVFTLTUTQKH	240	
Db	171	HPILGYPQDLSKLAFTSVYADLMDCGNKARTVAATNMNETSSRSFHVFTIVTQCH	230	
Qy	241	DEETKMDPEKAKISLVDLGASRATSTGATGARLKEGAEINRSISLTGLRVIAALADMS	300	
Db	231	DQLTGLDSEKYSKISLVLDGASRDSGARGMLKREGANKSLTTLGKVISALADMQS	290	
Qy	301	GKQKNQLVPRDSVLTWLLKDSLGGNSMTAAIATSPADINFEETLSTLRVADSAK	357	

Db 309 EATHINLSUTTLGRVIDVLADMATKAKAQSVAPFRDLSKLTFLKDSLGNSKTFMIAT 368
QY 336 ISPADINFEETSLTRYADSAK 357
Db 369 VSPSALNYEETSLTRYASRAR 390

RESULT 4

US-08-428-414A-3
; Sequence 3, Application US/08428414A
; Patent No. 5912166
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TITLE OF INVENTION: LEISHMANIASIS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428.414A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kadlecek, Ann T.
; REGISTRATION NUMBER: 39,244
; REFERENCE/DOCKET NUMBER: 210121.407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANDBERRY
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-428-414A-3

Query Match 38.5%; Score 705.5; DB 2; Length 955;
Best Local Similarity 42.7%; Pred. No. 6.9e-64;
Matches 163; Conservative 58; Mismatches 126; Indels 35; Gaps 8;
QY 7 IKVVYRVRFNARE--IDRGAKCIVRMENQOTILTPP-----PGAEKARKSGKTI 57
Db 13 VKVSVRVPLNERENNAPSGTKVTVAAKQAAVTVKVLGGNSNGAASMGTAARRA 72
QY 58 GKPAFADRSYWSF---DKNAPNVARQEDLFQDLGVPLLDNAPKGYNNCIFAYG 114
Db 73 ----FQDHVFWSVETPDACGATPATQADVFTIGYPLVQHAFDFGNSCLFAYG 128
QY 115 SYSMMG-----YKKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYN 168
Db 129 TYTMGADVSAISGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEYIN 188
QY 169 NPSTKG-----NLKVRHPSTGTPVEDLAKLVRSFOEINLMDGKNKARTVA 220
Db 189 GKRRKGVKGGEVYVDVREHPSRGVFLQGRVLEVGSLDDVRLIEIGNVRHTA 248
QY 221 NETSSRSRAVFTLTL----TQKHDEETKMDTEKAKISLVLDLAGSERATSTG 276
Db 249 NDRSSRSRAIIMLLREERTMTTKSETIRTAGKSSRMNLVLDLAGSERVAQS 308

QY 277 EGAEINRSLSLTGRVIAALADMSS-GKOKKNQLVPRDSVJTWLLKDSLGNSMTA 335
Db 309 EATHINLSUTTLGRVIDVLADMATKAKAQSVAPPROSKLTFLKDSLGNSKTF 368
QY 336 ISPADINFEETSLTRYADSAK 357
Db 369 VSPSALNYEETSLTRYASRAR 390

RESULT 5

PCT-US94-00324-1
; Sequence 1, Application PC/TUS9400324
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word, version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00324
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/006, 676
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 5004-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-00324-1

Query Match 38.5%; Score 705.5; DB 5; Length 955;
Best Local Similarity 42.7%; Pred. No. 6.9e-64;
Matches 163; Conservative 58; Mismatches 126; Indels 35; Gaps 8;
QY 7 IKVVYRVRFNARE--IDRGAKCIVRMENQOTILTPP-----PGAEKARKSGKTI 57
Db 13 VKVSVRVPLNERENNAPSGTKVTVAAKQAAVTVKVLGGNSNGAASMGTAARRA 72
QY 58 GKPAFADRSYWSF---DKNAPNVARQEDLFQDLGVPLLDNAPKGYNNCIFAYG 114
Db 73 ----FQDHVFWSVETPDACGATPATQADVFTIGYPLVQHAFDFGNSCLFAYG 128
QY 115 SYSMMG-----YKKEHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYN 168
Db 129 TYTMGADVSAISGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEYIN 188
QY 169 NPSTKG-----NLKVRHPSTGTPVEDLAKLVRSFOEINLMDGKNKARTVA 220
Db 189 GKRRKGVKGGEVYVDVREHPSRGVFLQGRVLEVGSLDDVRLIEIGNVRHTA 248
QY 221 NETSSRSRAVFTLTL----TQKHDEETKMDTEKAKISLVLDLAGSERATSTG 276

Db 249 NDRSRSHAIIMLLREERTTTKSGETIRTAGKSSRMNLVDLAGSERVAOSQVEGOQFK 308
QY 277 EGAEINRSLSTGRVIAALADMSS-GKQKKNQVPRDSVLTWLLKDSLGGSMTAMTAA 335
Db 309 EATHINLSLTTLGRVIDVLADMATKAKAQYSAVPRFSRUKTFLIKDSLGGSNKTFTMIAT 368
QY 336 ISPADINFEETLSTLYADSAK 357
Db 369 VSPSALNYEETLSTLYASRAR 390
RESULT 6
US-08-713-815A-4
; Sequence 4, Application US/08/713815A
; Patent No. 5830659
; GENERAL INFORMATION:
; APPLICANT: Russell J. Stewart
; TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED
; TITLE OF INVENTION: SEPARATIONS BY KINESINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.
; CITY: Sandy
; STATE: Utah
; COUNTRY: USA
; ZIP: 84070
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: AST Ascentia 900N
; OPERATING SYSTEM: DOS 6.22
; SOFTWARE: Word Perfect 6.0
; CURRENT APPLICATION DATA:
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Alan J. Howarth
; REGISTRATION NUMBER: 36,553
; REFERENCE/DOCKET NUMBER: T3214/U-2202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801)566-6633
; TELEFAX: (801)566-0750
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
US-08-713-815A-4

Query Match 31.7%; Score 582; DB 2; Length 411;
Best Local Similarity 40.0%; Pred. No. 1e-51;
Matches 142; Conservative 59; Mismatches 116; Indels 38; Gaps 10;

QY 6 NIKVVVRVPPFNAREIDRGAKCIVRMEGNQITLTTPPGAEEKARKSGKTIIMDGPRAFAFD 65
Db 12 SIKVYCRFRPLNDSEKAGSFVVKF-----PNNVEENC-----ISTAGKVLYFD 56
QY 66 RSYWSEFNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQTGSGKSYSM---MGYG 122
Db 57 KVF-----KPN-ASQEKVYNEAAKSIYTDVLGYNGTIFAYGQTSSGKTHMEGVIGDS 109
QY 123 KEHGVIPRICODMFRINELQKDLCTVEVSYLEIYNERVDRLLNPNSTGNLKVREHP 182
Db 110 VKQGIIPRIVNDIFNHIYAM--EVNLEPHIKVSYVEIYMDKIRDLDDYS-KVNLVSHDEK 166
QY 183 STGPVVEDLAKLVRSFQEIENLMDGKNKARTVAATNNETSSRSRSHAVFTLTLLTKQWIDE 242
Db 110 VKQGIIPRIVNDIFNHIYAM--EVNLEPHIKVSYVEIYMDKIRDLDDYS-KVNLVSHDEK 166
QY 183 STGPVVEDLAKLVRSFQEIENLMDGKNKARTVAATNNETSSRSRSHAVFTLTLLTKQWIDE 242
Db 167 NRVPYVKGATERFVSSPEDVPEVIEEGKSNRHIATVNNNEHSSRSRSHVFLINVQENLEN 226
QY 243 ETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302
Db 167 NRVPYVKGATERFVSSPEDVPEVIEEGKSNRHIATVNNNEHSSRSRSHVFLINVQENLEN 226
QY 243 ETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302

Db 227 QKKLS-----GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN 279
QY 303 QKKNQVPRDSVLTWLLKDSLGGSMTAMTAAISPADINFEETLSTLYADSAK 357
Db 280 KTH-----IPYRDSKLTROILOESLGGNARTTIVICCSPASFNESETKTLDFGRRAK 331
RESULT 7
US-08-713-815A-3
; Sequence 3, Application US/08/713815A
; Patent No. 5830659
; GENERAL INFORMATION:
; APPLICANT: Russell J. Stewart
; TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED
; TITLE OF INVENTION: SEPARATIONS BY KINESINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.
; CITY: Sandy
; STATE: Utah
; COUNTRY: USA
; ZIP: 84070
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: AST Ascentia 900N
; OPERATING SYSTEM: DOS 6.22
; SOFTWARE: Word Perfect 6.0
; CURRENT APPLICATION DATA:
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Alan J. Howarth
; REGISTRATION NUMBER: 36,553
; REFERENCE/DOCKET NUMBER: T3214/U-2202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801)566-6633
; TELEFAX: (801)566-0750
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
US-08-713-815A-3

Query Match 31.7%; Score 582; DB 2; Length 441;
Best Local Similarity 40.0%; Pred. No. 1.1e-51;
Matches 142; Conservative 59; Mismatches 116; Indels 38; Gaps 10;

QY 6 NIKVVVRVPPFNAREIDRGAKCIVRMEGNQITLTTPPGAEEKARKSGKTIIMDGPRAFAFD 65
Db 12 SIKVYCRFRPLNDSEKAGSFVVKF-----PNNVEENC-----ISTAGKVLYFD 56
QY 66 RSYWSEFNAPYARQEDLFQDLGVPLLDNAPKGYNNCIFAYGQTGSGKSYSM---MGYG 122
Db 57 KVF-----KPN-ASQEKVYNEAAKSIYTDVLGYNGTIFAYGQTSSGKTHMEGVIGDS 109
QY 123 KEHGVIPRICODMFRINELQKDLCTVEVSYLEIYNERVDRLLNPNSTGNLKVREHP 182
Db 110 VKQGIIPRIVNDIFNHIYAM--EVNLEPHIKVSYVEIYMDKIRDLDDYS-KVNLVSHDEK 166
QY 183 STGPVVEDLAKLVRSFQEIENLMDGKNKARTVAATNNETSSRSRSHAVFTLTLLTKQWIDE 242
Db 167 NRVPYVKGATERFVSSPEDVPEVIEEGKSNRHIATVNNNEHSSRSRSHVFLINVQENLEN 226
QY 243 ETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302
Db 227 QKKLS-----GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN 279
QY 303 QKKNQVPRDSVLTWLLKDSLGGSMTAMTAAISPADINFEETLSTLYADSAK 357

ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50451
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-097-767A-26

Query Match 5.0%; Score 92; DB 4; Length 402;
Best Local Similarity 22.4%; Pred. No. 0.41;
Matches 77; Conservative 43; Mismatches 127; Indels 96; Gaps 19;

QY 35 QTLTPPPGAE-----KARKSGKTIWDGPKAFADRSY-----WSFDKNAP 76
DB 60 QPVLPPEKKVPFELWYTELKRTSSITL-----ATRMNLYLVGRTGGVWMEFGKDG- 113
QY 77 NYAQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYMMG-YGKEHGVIPIRICODM 135
DB 114 -----DTHLLGDNPRLGFG-----GRYQDLIGNKGLTYSMGR--AEM 150
QY 136 FRINELQKKNLTCTVEVSYLEIYNERNVRLDLPSTKGNLKVREHPSTGPVEDLAKLV 195
DB 151 TRAVNDLAKKKAASATLK-----KKVKMQMPKAAKLAADPQADT-KSKLVKLV 202
QY 196 VRSFOEINLMDGKARTVAATNMNETSSRSYAVFTLT-----QKWHDETRMDTEK 250
DB 203 V-----MVCGRFNTVSR-TVDAGFNSOHGV-TLTVTGKQVQKW-----DR 243
QY 251 VAKISVLDSLGRATSGATGARLKEGAEINRSSTLGRVIAALADSSGKQKKNQVLP 310
DB 244 ISKAAF-EWADHTAVIPDMQKLGKIDKNEARIIVALKVNTTAAAD-ASAKNKKKSKP 301
QY 311 YRDSVLTWLLKSLGG-----NSWTAM-TAAISPADINTEE 345
DB 302 IGDDV-----PILGGIINVQINSVLVFDQLARFAVQDHNKKE 338

RESULT 11
US-08-732-433-1
Sequence 1, Application US/08732433
Patent No. 6063611
GENERAL INFORMATION:
APPLICANT: van Solingen, Dieter
TITLE OF INVENTION: No. 6063611al Alkaline Cellulase and
TITLE OF INVENTION: Method of Producing Same
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,433
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05651
FILING DATE: 26-APR-1996
ATTORNEY/AGENT INFORMATION:

NAME: Stone, Christopher L.
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC287-US2
TELEPHONE: 650-846-7620
TELEFAX: 650-846-6504
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-732-433-1

Query Match 5.0%; Score 92; DB 3; Length 574;
Best Local Similarity 19.5%; Pred. No. 0.72;
Matches 68; Conservative 43; Mismatches 134; Indels 104; Gaps 13;

QY 28 IVRMENGTILTP-----PPGAEEKARKSGKTI--MDGPKAFADRSY--WSFDKNAP 76
DB 202 IVRQGGQNDIRPLVLPMTATSQPLNLLNYQTIDKDDPNLIATVHYGFWPFSVNIA 261
QY 77 NYAR-QEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYMMGYGKEHGVIP-----R 130
DB 262 GYTREEDSKRE-----IIEFTDRVHHTFVARGIPVVLGFEGLLGFDKHTGVIQOGEK 316
QY 131 IQDMFRINELQKKNLTCTVEVSYL-----EYNERVRLDLPSTKGNLKV 177
DB 317 FFYELIHLNER-----DITHMLWDNQHFNRHTYEWYDELFQDLARFAVQDHNKKE 366
QY 178 VRE-----HPSGTGPVEDLAKLVVRSFOEINLMDGKARTVAATNMNETSSRSYAVFTL 233
DB 367 VAESNFIYKQGDRIADAT-----VTQLHGNETLGLQANGQRLTPQGDYELN 416
QY 234 TITQKWHDEETKMDTEKVKAKISVLDSLGRATSGATGARLKEGAEINRSSTLGRVIA 293
DB 417 RLTVKAH-----VLSAIGSGTLGTNGMVTAEFNRCADWHFRYNT-----456
QY 294 ALADSSGKQKKNQVLPYRDSVL-----TWLLKDSLGSGNSMTAMIA 334
DB 457 -----YRTPVQLSTQGHVSFNFSIPASFNGNSLATMEA 488

RESULT 12
US-08-368-281-2
Sequence 2, Application US/08368281
Patent No. 5721113
GENERAL INFORMATION:
APPLICANT: Libermann, Towia A
APPLICANT: Oettgen, Joerg P
APPLICANT: Kunsch, Charles A
TITLE OF INVENTION: NERF Genes
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,281
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833

TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 521 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-368-281-2

Query Match 5.0%; Score 91; DB 1; Length 521;
Best Local Similarity 19.9%; Pred. No. 0.78;
Matches 83; Conservative 61; Mismatches 145; Indels 128; Gaps 21;

QY 39 TTPPGAEEKARK-SGKTIMDGPKAFADRSYWSF-----DKN-APNY-----ARQEDLFQ 86
DB 119 SPELGIKKKPREGKNTTY-----LWFEFLDLLQDKNTCPRYIKWTQREKGIK 167
QY 87 DLGVPLLDNAFKYNNICIFAYGOTGSGKSYSMWG-----YKHEGVIPRI-----COD 134
DB 168 -----LVDS--KAVSK-LWGKHKNNPDNMYETMGRALRYYYQGLAKVGEQRLVYQFKD 219
QY 135 MFRINELQDKNLTCT-----TVEVSYLEIYNERVRLDLPST-----172
DB 220 MPKNIVVDDDKSETCNEGLAGTTDEKSLERVSLSAESLLKAASSVRSKGNSSPINCRA 279
QY 173 -KGNLKV-----RHPSTGPVVEDLAKLVRSFOEINLWDE--GNKARTVAAT 218
DB 280 EKVARKVNIYSPCHDASSRPTTASVSATAAPRTVRVAMQVPMVPMVSLGQKISTVAVQ 339
QY 219 NMN-----ETSSRSYHAFVTLT-----TQKWHDEETKMDTKVAKISLVLDLAGSE 263
DB 340 SVNAGAPLITSTPTTATSKVVIQTIPTVMPASTENGDKITMQPAKIITIPATOLAQCQ 399
QY 264 RATSTGATGA-----RLKEGAENR-----SLSTLGRVIAALAD 297
DB 400 LOTKSNLTGSGSINIVCTPLAVRALTPVSIANGTPVRLSMPTQOASQTPPRVISAVI- 458
QY 298 MSSGQKKNQVLPRD--SVLTWLL---KDSLGNSMTAMIAISPADINFEETLST 349
DB 459 --KGAEVKSEAVAKQEDHVKTLLELVEEKPADGNKKTVTHVVVVSAPSALPVTMKT 513

RESULT 13

US-08-368-281-4
Sequence 4, Application US/08368281
Patent No. 572113
GENERAL INFORMATION:
APPLICANT: Libermann, Towia A
APPLICANT: Oetgen, Joerg P
APPLICANT: Kunsch, Charles A
TITLE OF INVENTION: NERF Genes
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,281
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 581 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-368-281-4

Query Match 5.0%; Score 91; DB 1; Length 581;
Best Local Similarity 19.9%; Pred. No. 0.94;
Matches 83; Conservative 61; Mismatches 145; Indels 128; Gaps 21;

QY 39 TTPPGAEEKARK-SGKTIMDGPKAFADRSYWSF-----DKN-APNY-----ARQEDLFQ 86
DB 179 SPELGIKKKPREGKNTTY-----LWFEFLDLLQDKNTCPRYIKWTQREKGIK 227
QY 87 DLGVPLLDNAFKYNNICIFAYGOTGSGKSYSMWG-----YKHEGVIPRI-----COD 134
DB 228 -----LVDS--KAVSK-LWGKHKNNPDNMYETMGRALRYYYQGLAKVGEQRLVYQFKD 279
QY 135 MFRINELQDKNLTCT-----TVEVSYLEIYNERVRLDLPST-----172
DB 280 MPKNIVVDDDKSETCNEGLAGTTDEKSLERVSLSAESLLKAASSVRSKGNSSPINCRA 339
QY 173 -KGNLKV-----RHPSTGPVVEDLAKLVRSFOEINLWDE--GNKARTVAAT 218
DB 340 EKVARKVNIYSPCHDASSRPTTASVSATAAPRTVRVAMQVPMVPMVSLGQKISTVAVQ 399
QY 219 NMN-----ETSSRSYHAFVTLT-----TQKWHDEETKMDTKVAKISLVLDLAGSE 263
DB 400 SVNAGAPLITSTPTTATSKVVIQTIPTVMPASTENGDKITMQPAKIITIPATOLAQCQ 459
QY 264 RATSTGATGA-----RLKEGAENR-----SLSTLGRVIAALAD 297
DB 460 LOTKSNLTGSGSINIVCTPLAVRALTPVSIANGTPVRLSMPTQOASQTPPRVISAVI- 518
QY 298 MSSGQKKNQVLPRD--SVLTWLL---KDSLGNSMTAMIAISPADINFEETLST 349
DB 519 --KGAEVKSEAVAKQEDHVKTLLELVEEKPADGNKKTVTHVVVVSAPSALPVTMKT 573

RESULT 14

US-07-777-715-7
Sequence 7, Application US/07777715
Patent No. 5273889
GENERAL INFORMATION:
APPLICANT: Potter, Andrew
APPLICANT: Campos, Manuel
APPLICANT: Hughes, Huw P. A.
TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/777,715
FILING DATE: 19911016
CLASSIFICATION: 424

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 29310-2001320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-777-715-7

Query Match      4.9%  Score 90; DB 1; Length 1098;
Best Local Similarity 20.3%  Pred. No. 3.4;
Matches 85; Conservative 70; Mismatches 167; Indels 96; Gaps 21;

Qy  7 IKVVVRVP---FNAREI-----DRGAKCIVRMEGNQITLTPPGAAEERKARKSGKTMIDGP 59
   ::::: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  74 LEEVLNAPSKNLNPRIKDSMDNIKRIIVLELQSGSETRFT---CEYDDATVNAVEFLN-- 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  60 KAFAPDRSYWFDKNAPYARQEDLFQDLGVPLLDNAFKGYNNCIFA--YQGTSGSKSYS 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  129 KWIIFCQSIYS-----TMTGDLSPRLTTLNGLKNTLTATKSLGHLKAGOSLT 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  118 MWGYCKEHGV-----IPRICQDMFRRINELOKDKNLTCTVEVSYLEIYNRVRDLLNPS 171
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  177 QAGSLTKGAKKIILYIPQYDTEQGNGLQ---DLVKAEEELGIEVQREERNNIATAQ 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  172 TK-GNLKV-----REHPSTGPYVEDIAKLVV--RSFOEINLMDENKARTV----- 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  234 TSLGTIQTGALTEGIVLSAPQIDKLLQKTGAGLSDVIGSLSGATAALVADKNASTAKKVGA 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  306 -----NQLV-PYRDSVLTWLKDSL-GGNSMTAMIAAISPADINPEETLSTLRYADSA 356
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  408 GFELANQVVGNIITKAVSSYIILQAVAAAGLSSTGPPVAALIASTVSL--AISPLAFAGIA 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-08-170-126-2
; Sequence 2, Application US/08170126
; Patent No. 5594107
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW
; APPLICANT: CAMPOS, MANUEL
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,126
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,715
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/571,301
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0013.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-170-126-2

Query Match      4.9%  Score 90; DB 1; Length 1098;
Best Local Similarity 20.3%  Pred. No. 3.4;
Matches 85; Conservative 70; Mismatches 167; Indels 96; Gaps 21;

Qy  7 IKVVVRVP---FNAREI-----DRGAKCIVRMEGNQITLTPPGAAEERKARKSGKTMIDGP 59
   ::::: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  74 LEEVLNAPSKNLNPRIKDSMDNIKRIIVLELQSGSETRFT---CEYDDATVNAVEFLN-- 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  60 KAFAPDRSYWFDKNAPYARQEDLFQDLGVPLLDNAFKGYNNCIFA--YQGTSGSKSYS 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  129 KWIIFCQSIYS-----TMTGDLSPRLTTLNGLKNTLTATKSLGHLKAGOSLT 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  118 MWGYCKEHGV-----IPRICQDMFRRINELOKDKNLTCTVEVSYLEIYNRVRDLLNPS 171
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  177 QAGSLTKGAKKIILYIPQYDTEQGNGLQ---DLVKAEEELGIEVQREERNNIATAQ 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  172 TK-GNLKV-----REHPSTGPYVEDIAKLVV--RSFOEINLMDENKARTV----- 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  234 TSLGTIQTGALTEGIVLSAPQIDKLLQKTGAGLSDVIGSLSGATAALVADKNASTAKKVGA 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  216 -----AATNMNE-----TSRSRAVFTLTITQKHDEETKMDTEKVA-KISLVDLAGS----- 262
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  294 LGSVLGMDLDEALQNNNSQHALAKAGL-----ELTNSLIENIANSYKTLDEFGEQISQ 347
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  263 -----ERATSTGATGARLKE-----GAEINRSLSTLGRVIAAALADMSSGKOKK--- 305
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  348 FGSKLQNTKGLTGLDKLKNIGGLDKAGLDVIGSLSGATAALVADKNASTAKKVGA 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  306 -----NQLV-PYRDSVLTWLKDSL-GGNSMTAMIAAISPADINPEETLSTLRYADSA 356
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  408 GFELANQVVGNIITKAVSSYIILQAVAAAGLSSTGPPVAALIASTVSL--AISPLAFAGIA 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: April 25, 2001, 10:13:24
Job time: 240 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:14:07 ; Search time 46.78 Seconds
(without alignments)
524.457 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357
Perfect score: 1834
Sequence: 1 MSGGNIKVVVRPFNARE.....PADINFEETLSTRYADSAK 357

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_67.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1147	62.5	1695	2 A56921	kinesin family prote
2	1106.5	60.3	1584	1 JN0114	kinesin-related pr
3	1106.5	60.3	1584	2 T15822	kinesin-like prote
4	1106	60.3	1150	1 A55289	kinesin-like prote
5	1028	56.1	1921	2 T13827	kinesin-73 - fruit
6	965.5	52.6	1576	2 T29237	hypothetical prote
7	878.5	47.9	928	2 T16759	hypothetical prote
8	793.5	43.3	1121	2 T13750	kinesin-like prote
9	792.5	43.2	1121	2 T13796	kinesin-related pr
10	743.5	40.5	699	1 S38982	kinesin-related pr
11	712	38.8	742	1 S58691	kinesin-related pr
12	705.5	38.5	955	2 A47334	LeKin kinesin-rela
13	701	38.2	747	1 A57107	kinesin-related pr
14	694	37.8	786	2 A53939	kinesin homolog KH
15	689	37.6	701	1 B44259	kinesin-related pr
16	673	36.7	1229	2 T48959	kinesin-like prote
17	669	36.5	928	2 T10164	kinesin heavy chai
18	666.5	36.3	1231	2 A54803	microtubule-associ
19	663	36.2	332	2 C48835	kinesin-like prote
20	651	35.5	784	1 A55236	kinesin-related pr
21	649	35.4	793	2 JC5831	kinesin-related pr
22	646	35.2	929	2 T19132	kinesin [imported]
23	643	35.1	1226	2 I51617	kinesin-like prote
24	641	35.0	1295	2 A56514	chromokinesin - ch
25	634	34.6	963	1 A41919	kinesin heavy chai
26	633.5	34.5	935	2 T51930	kinesin [imported]
27	632	34.5	881	2 T84737	kinesin heavy chai
28	629.5	34.3	294	2 S38983	kinesin-related pr
29	627	34.2	672	2 S54351	kinesin oam-3 - Ca

30 615.5 33.6 2954 2 T14156 kinesin-related pr
31 612 33.4 1056 1 G02157 kinesin-like spind
32 611.5 33.3 1032 2 I38510 neuronal kinesin h
33 610 33.3 968 2 T51933 kinesin motor prot
34 603 32.9 1075 2 T45570 kinesin-like prote
35 601 32.8 1027 2 S37711 kinesin heavy chai
36 599 32.7 2663 1 S28261 centromere protein
37 596.5 32.5 967 1 A35075 kinesin heavy chai
38 595.5 32.5 1130 2 T21134 hypothetical prote
39 592 32.3 1031 1 A38713 kinesin heavy chai
40 590 32.2 1254 2 T18277 kinesin heavy chai
41 582 31.7 975 1 A31497 kinesin heavy chai
42 579 31.6 1067 2 S33417 kinesin-like prote
43 572 31.2 1060 1 A40264 kinesin-related pr
44 568.5 31.0 843 2 S44868 kinesin heavy chai
45 568.5 31.0 883 2 T40128 kinesin-like prote

ALIGNMENTS

RESULT 1
A56921
kinesin family protein KIF1a - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: A56921
R:Okada, Y.; Yamazaki, H.; Sekine-Aizawa, Y.; Hirokawa, N.
Cell 81, 769-780, 1995
A:Title: The neuron-specific kinesin superfamily protein KIF1a is a unique monomeric
A:Reference number: A56921; MUID:95292344
A:Accession: A56921
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1695 <RES>
A:Cross-references: GB:ID29951; NID:g976234; PID:BAA06221.1; PID:g976235
C:Superfamily: kinesin-related protein unc-104; kinesin motor domain homology; plecks
C:Keywords: P-loop
F:6-360/Domain: kinesin motor domain homology <KNOT>
F:97-104/Region: nucleotide-binding motif A (P-loop)

Query Match	62.5%	Score 1147;	DB 2;	Length 1695;
Best Local Similarity	61.8%	Pred. No. 6.9e-78;		
Matches 225;	Conservative 58;	Mismatches 57;	Indels 24;	Gaps 6;
Qy	4	GNIKVVVRPFNAREIDRGAKCIVRMENQOTILTPPGAEKARKSGKTMDGPKAFA	63	
Db	3	GASVAVVRPFNSREMSRDSKCIQMSGSTTTIVNPKPKET-----PKSFS	51	
Qy	64	FRSWSFDKNAP---NYARQEDLFQDLGVPLDPAFGYNNCIFYGQTGSGKSYSMWG	120	
Db	52	FDYSYWS--HTSPEDINYSQKVYRDIQGEMLQHAFFGVNCIFYGQTGAGKSYTMWG	109	
Qy	121	YGK--EHGVIPTRICQDMFRINELOKQKNTCTVEYSLEYINERVRDLNPFSTGKNLKV	178	
Db	110	KQKQDQGLIPOLCEDLFSRINDTND--NMSYSVEYSYMEIYCERVDLLNPKNGLRV	168	
Qy	179	REHPTGPGYEDLAKLVVSFOEINLMDGSKARVATNNTSSRSRSHAVFTLTQK	238	
Db	169	REHPLGPGYEDLAKLVVSFOEINLMDGSKARVATNNTSSRSRSHAVFTLTQK	228	
Qy	239	WHEETKMDTEKVAKISLVDLGASERATSTGATGARLKEGAEINRSLSLTGRVIAALADM	298	
Db	229	RHDAETNITTEKYSKISLVDLGASERADSTGAKGTRLEKAGINRSLSLTGRVIAALAE	288	
Qy	299	SSG-----KQKKQLVPYSDSVLWLLKDSIGNSMTAMIAISPADINFEETLSTLYA	353	
Db	289	DSGPNKKKKKKTKDFIPYRDSVLTWLLRENLGNSTRTAWAALSPADINDETSTLYA	348	
Qy	354	DSAK 357		
Db	349	DRAK 352		

A: Introns: 34/1; 84/2; 140/3; 200/2; 285/3; 339/2; 490/2; 580/1; 631/3; 672/1; 738/2;
C: Superfamily: kinesin-related protein unc-104; kinesin motor domain homology; plecks

RESULT 2
JN0114
kinesin-related protein unc-104 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 29-May-1998
C: Accession: JN0114
R: Otsuka, A.J.; Jeyaparakash, A.; Garcia-Anoveros, J.; Tang, L.Z.; Fisk, G.; Hartshorne,
Neuron 6, 113-122, 1991
A: Title: The C. elegans unc-104 gene encodes a putative kinesin heavy chain-like protein
A: Reference number: JN0114; MUID:91097805
A: Accession: JN0114
A: Molecule type: mRNA
A: Residues: 1-1584 <OTS>
A: Cross-references: GB:M58282
A: Note: 598-Thr and 930-Met were also found
C: Genetics:
A: Gene: unc-104
C: Superfamily: kinesin-related protein unc-104; kinesin motor domain homology; pleckstrin
C: Keywords: ATP; microtubule binding; P-loop
F: 4-353/Domain: kinesin motor domain homology <KMOT>
F: 93-100/Region: nucleotide-binding motif A (P-loop)
F: 1285-1287/Region: cell attachment (R-G-D) motif
F: 99/Binding site: ATP (Lys) #status predicted

Query Match 60.3%; Score 1106.5; DB 1; Length 1584;
Best Local Similarity 61.2%; Pred. No. 7e-75;
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

Qy 6 NIKVVVRYPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIMDGPKAFAD 65
Db 3 SVKAVRVRPNQREISNTSKCVLVNGNTT-----NGHSINKENSFNFD 50
Qy 66 RSYWFDKNAPYARQEDLFODLVPLLDNAPKGYNNCIFAYGQTGSGKSYMMGYG-- 123
Db 51 HSYWSEFARNDPHFTQKQVYELGVEMLEHAPEGYNVCIFAYGQTGSGKSYMMGKANDP 110
Qy 124 -EHGVIPRICODMFRINELQKMLCTVEVSYLEIYNERVRLDNLNPNSTGNLKVREHP 182
Db 111 DEMGIIPRLCNDLFARIDN--NNDKDVQYSVEVSYMEIYCEYKVKDLLNPNSGGNLVRREHP 169
Qy 183 STGPVVEDLAKLVRSFOEINLMDGKARTVAATNNSTSSSRSHAVFTLTQKHDE 242
Db 170 LLGPTVDDLTMAVCSYHDICNLMDGKARTVAATNNSTSSSRSHAVFTLTQKHCA 229
Qy 243 ETAMDTEKVKAKISLVDLGSRATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302
Db 290 KSKNKGVIPIRDSVLTLLRENLGNSKTAMLAALSPADINFDLTSLTLRYADRAK 345
Qy 303 OKKNO-LVPIRDSVLTLLKSLGNSMTAMTAAISPADINFEETLTLRYADSAK 357
Db 290 KSKNKGVIPIRDSVLTLLRENLGNSKTAMLAALSPADINFDLTSLTLRYADRAK 345

RESULT 3
JN15822
kinesin-like protein unc-104 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-Jan-2000
C: Accession: J15822
R: Du, Z.

submitted to the EMBL Data Library, February 1996
A: Description: The sequence of C. elegans cosmid C52E12.
A: Reference number: Z18412
A: Accession: J15822
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1584 <DUZ>
A: Cross-references: EMBL:U05135; NID:g1208884; PID:g1208885; PIDN:AAA93453.1; CESP:unc-1
C: Genetics:
A: Gene: CESP:unc-104

Query Match 60.3%; Score 1106.5; DB 2; Length 1584;
Best Local Similarity 61.2%; Pred. No. 7e-75;
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

Qy 6 NIKVVVRYPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIMDGPKAFAD 65
Db 3 SVKAVRVRPNQREISNTSKCVLVNGNTT-----NGHSINKENSFNFD 50
Qy 66 RSYWFDKNAPYARQEDLFODLVPLLDNAPKGYNNCIFAYGQTGSGKSYMMGYG-- 123
Db 51 HSYWSEFARNDPHFTQKQVYELGVEMLEHAPEGYNVCIFAYGQTGSGKSYMMGKANDP 110
Qy 124 -EHGVIPRICODMFRINELQKMLCTVEVSYLEIYNERVRLDNLNPNSTGNLKVREHP 182
Db 111 DEMGIIPRLCNDLFARIDN--NNDKDVQYSVEVSYMEIYCEYKVKDLLNPNSGGNLVRREHP 169
Qy 183 STGPVVEDLAKLVRSFOEINLMDGKARTVAATNNSTSSSRSHAVFTLTQKHDE 242
Db 170 LLGPTVDDLTMAVCSYHDICNLMDGKARTVAATNNSTSSSRSHAVFTLTQKHCA 229
Qy 243 ETAMDTEKVKAKISLVDLGSRATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302
Db 290 KSKNKGVIPIRDSVLTLLRENLGNSKTAMLAALSPADINFDLTSLTLRYADRAK 345
Qy 303 OKKNO-LVPIRDSVLTLLKSLGNSMTAMTAAISPADINFEETLTLRYADSAK 357
Db 290 KSKNKGVIPIRDSVLTLLRENLGNSKTAMLAALSPADINFDLTSLTLRYADRAK 345

RESULT 4
A55289
kinesin-like protein KIF1B - mouse
C: Species: Mus musculus (house mouse)
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C: Accession: A55289
R: Nangaku, M.; Sato-Yoshitake, R.; Okada, Y.; Noda, Y.; Takemura, R.; Yamazaki, H.; H
Cell 79, 1209-1220, 1994
A: Title: KIF1B, a novel microtubule plus end-directed monomeric motor protein for tra
A: Reference number: A55289; MUID:95094296
A: Accession: A55289
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-1150 <NAN>
A: Cross-references: GB:D17577; NID:g407338; PIDN:BAA04503.1; PID:g407339
C: Superfamily: kinesin-related protein KIF1B; kinesin motor domain homology
C: Keywords: P-loop
F: 6-354/Domain: kinesin motor domain homology <KMOT>
F: 97-104/Region: nucleotide-binding motif A (P-loop)

Query Match 60.3%; Score 1106; DB 1; Length 1150;
Best Local Similarity 61.18; Pred. No. 4.8e-75;
Matches 218; Conservative 59; Mismatches 64; Indels 16; Gaps 5;

Qy 4 GNIKVVVRYPFNAREIDRGAKCIVRMEGNOTILTPPPGAEEKARKSGKTIMDGPKAFAD 63
Db 3 GASVAVRVRPNFNSRSTSKESKCIIOGNSSTSIINPNKE-----APKSFS 51
Qy 64 FDRSYWSF-DKNAPYARQEDLFODLVPLLDNAPKGYNNCIFAYGQTGSGKSYMMGYG 122
Db 52 FDSYWSHTSPEDPCFASQNFVNDIGKEMLLHAPEGYNVCIFAYGQTGAGKSYTMGKQ 111
Qy 123 KEHG--VTPRICODMFRINELQKMLCTVEVSYLEIYNERVRLDNLNPNSTGNLKVRE 180
Db 112 BESQAVIIPQLCEELFEKIND--NCNEEMSYSEVSYMEIYCEYKVKDLLNPNKGNLRYRE 170
Qy 181 HPSGTGPYVEDLAKLVRSFOEINLMDGKARTVAATNNSTSSSRSHAVFTLTQKH 240
Db 171 HPLGPGYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNNSTSSSRSHAVFTVTRKQ 230

[illegible]

Query Match	56.1%	Score 1028;	DB 2;	Length 1921;
Best Local Similarity	60.3%;	Pred. No.	7.4e-69;	
Matches 213;	Conservative 45;	Mismatches 85;	Indels 10;	Gaps 4;

Qy	7	IKVVVRFPNAREIDRGAKCIVRMGEHOTILPPGAEERKARKSGKTIMDGPAFAFDR	66
Db	6	IKVAVRFPNAREIETDKCIVEMEKOOTLQNQPPELKERKQ-----PKTFAPDH	58
Qy	67	SYWFOKNAPYARQEDFDOLGVPLLDNAFKYNNCFAYGTGSGKYSMMGYGKEHG	126
Db	59	CIFYSLNPEDENPASQGVTFDCVGRGILDONAFQYNACIFAYGTGSGKSYTMNGTQESKG	118
Qy	127	VIPRICODMFRINELQOKNKUICTVFVSYLEITYNERVDLLNPS--TKGNLKVREHPSTG	185
Db	119	IIPRLCDQLFSIAIN-KSTPELMKYVEVSEMYNEIKYKHVDLLDPKPNKOSLARHEHVMG	177
Qy	186	PYVEDIAKLIVRSFOETENIMDEGHNKARTVAATNMNETSSBSHAVFTLLTLTKQWHDET	245
Db	178	PYVDGLSQLANTVSQDIDLMLTEGNNKSTVAATNMNAESSKSHAVSVVLITQTLDQATG	237
Qy	246	MDTEKVAKISLVDLAGSERATSTGARLUKEGAENIRLSLTGLRVTAALADMSMSGKKK	305
Db	238	VSGEKSRSMLVDLAGSERAVTGVAGDRLKEGSINKSLTTGLVLISKLAQSNKKSG	297
Qy	306	N-QLVYPYRDVLTWLLKDSLGNSWTAMIAAISPADINFETLTSTRYADSAK	357
Db	298	NDFVPYRDVLTWLLKDNIGNSSTYMWATISPSADNYEBTSLTRYADR	350

RESULT 6
T929237
hypothetical protein F56E3.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T929237
R:Gatung, S.; Wu, X.
submitted to the EMBL data Library, November 1995
A:Description: The sequence of C. elegans cosmid F56E3.
A:Reference number: Z20592
A:Accession: T929237
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1576 <ANT>
A:Cross-references: EMBL:U41536; PIN:AAH52613.1; GSPDB:GN00028; CESP:F56E3.3

[illegible]

```

Query Match      47.9%; Score 878.5; DB 2; Length 928;
Best Local Similarity 51.1%; Pred. No. 4.6e-58;
Matches 189; Conservative 54; Mismatches 100; Indels 27; Gaps 6;

QY 1 MSGGGNKKVVVRVFNAEIDRGAKCIVRMGEQNQTILPTPPGAEAKRKSCKTMDGPK 60
Db 1 MKGDSIIIVAVRPPNDREKTRNCKLVIEPDEETVIRDPKTNDE-----K 48

QY 61 AFADRSTYSWFO-----KNR-----PNYARQEDLQDGLGVPLDNAFKGNYNCIFYAQ 109
Db 49 RFTYDHSYWSHGDFSEKKNGYLEPTDPIHADRQRRVEDLGRGVLANAGYNCISLIYAG 108

```


Qy	7	IKVVVRVRFNAREIDRGAKCIVRMENQTI--LTPPGAEEKARKSGKTI	MDGPKAF	64
		: : : : : : : :		
Dd	9	VKVVVRCRPMNSKETISQGHKRIIVEMDNKRLVEVTNPKGPPGPKNS	-----FTF	58

QY 336 ISPADINFEETLSTLYADSAK 357

QY 336 ISPADINFEETLSTLYADSAK 357

F:100-107/Region: nucleotide-binding motif A (P-loop)
F:369-599/Domain: helical rod #status predicted <ROD>
F:600-701/Domain: tail globular #status predicted <GTL>
F:106/Binding site: ATP (Lys) #status predicted

Query Match		37.6%	Score 689;	DB 1;	Length 701;
Best Local Similarity		45.6%	Pred. No. 5.6e-44;		
Matches 164;		Conservative 50;	Mismatches 108;	Indels 38;	Gaps 10;
Qy	6	NIKVVVRVPPNAREID	---RGACIVRMENGTILTPPPGAEEKARKSGKTINDGPKAF	62	
Db	14	NKVVVRCRPLNEKESMCRYQAVSVDMRGITIV	-----HKTSSN----	EPKTF	61
Qy	63	AFDRSYMSFDKNAPNYAQEDLFQDLGVPLLDNAFKYNNCIFAYGQTGSGKSYSMGYG	122		
Db	62	TEDTVF-----GPE-SKQLDVYNLTARPIIDSVLEGYNGTIFAYGQTGKTFTMEGVR	114		
Qy	123	KE---HGVIPLICODMPRINELQKDKNLCTVEVSYLEIYNERVROLNPNSTKGNLKV	179		
Db	115	AVPGLRGVIPSFAHIFGHIAREGDTRF--LVRVSIETIENEVRDLCKGQOTQRLVK	172		
Qy	180	EHPSTGPYVEDLAKLVYRSFQEIENLMDENKARTVAATNMNETSSRSRAVFTLTL--TQ	237		
Db	173	ERPDVGVIKDLISAYVYNADDMDRIMTLGHKNRSYVGATNMNEHSSRSRAIFTITIECSE	232		
Qy	238	KWDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSLTLGRVIAALAD	297		
Db	233	KGVDGNMHV---RMGKLHLVDLAGSERQAKTGATGORLKEATKINLSLTLGNVISALVD	289		
Qy	298	MSSGOKKNQLVPYRDSVLTWLLKDSLGNSMTAMIAISPADINFETLSTLRVADSAR	357		
Db	290	-----GKSTHVPYRNSKLTLLLODSLGGNSKTMCMCANIGPADYNYDETISTLRVANRAK	343		

Search completed: April 25, 2001, 10:14:12
Job time: 288 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:18:28 ; Search time 26.31 Seconds
(without alignments)
464.812 Million cell updates/sec

Title: US-09-235-416-1_COPY_1_357

Perfect score: 1834

Sequence: 1 MSGGNIKVVVRPFNARE.....PADINFEETLRLYADSAK 357

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	62.8	1690	1 KFLA_HUMAN	Q12756 homo sapien
2	1147	62.5	1695	1 KFLA_MOUSE	P33173 mus musculus
3	1124	61.3	1103	1 KFLC_HUMAN	Q43896 homo sapien
4	1119	61.0	1816	1 KFLB_HUMAN	O60333 homo sapien
5	1110	60.5	1816	1 KFLB_MOUSE	O60375 mus musculus
6	1106.5	60.3	1584	1 U104_CAEEL	P23678 caenorhabdi
7	1085.5	59.2	1097	1 KFLD_RAT	O35787 rattus norv
8	1073.5	58.5	699	1 KFLB_RAT	O88658 rattus norv
9	743.5	40.5	699	1 K122_STRPU	P46872 strongyloce
10	712	38.8	742	1 K121_STRPU	P46871 strongyloce
11	705.5	38.5	955	1 KINL_LEICH	P46865 leishmania
12	704	38.4	747	1 KFLB_HUMAN	Q15066 homo sapien
13	701	38.2	747	1 KFLB_MOUSE	O61771 mus musculus
14	694	37.8	786	1 FL10_CHLRE	P46869 chlamydomon
15	689	37.6	701	1 KFLA_MOUSE	P28741 mus musculus
16	684	37.3	702	1 KFLA_HUMAN	O9Y496 homo sapien
17	669	36.5	928	1 KINH_NEUCR	P48467 neurospora
18	666.5	36.3	1231	1 KFLA_MOUSE	P33174 mus musculus
19	657	35.8	1222	1 KFLA_HUMAN	O95339 homo sapien
20	656	35.8	796	1 KFLC_RAT	O55165 rattus norv
21	654	35.7	796	1 KFLC_MOUSE	O35066 mus musculus
22	651	35.5	784	1 K168_DROME	P46867 drosophila
23	649	35.4	793	1 KFLC_HUMAN	O14782 homo sapien
24	634	34.6	963	1 KINH_HUMAN	P33176 homo sapien
25	633.5	34.5	935	1 KINH_SYNRA	O43093 syncephalus
26	633	34.5	963	1 KINH_MOUSE	O61768 mus musculus
27	627	34.2	672	1 OSW3_CAEEL	P46873 caenorhabdi
28	612	33.4	1057	1 EGS_HUMAN	P52732 homo sapien
29	611.5	33.3	1032	1 KINH_HUMAN	Q12840 homo sapien
30	608.5	33.2	1027	1 KINN_MOUSE	P33175 mus musculus
31	604	32.9	957	1 KFLC_HUMAN	O60282 homo sapien
32	603	32.9	956	1 KFLC_MOUSE	P28738 mus musculus
33	599	32.7	2663	1 CENE_HUMAN	Q02224 homo sapien

RESULT 1
KFLA_HUMAN
ID KFLA_HUMAN STANDARD; PRT; 1690 AA.
AC Q12756;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1A (AXONAL TRANSPORTER OF SYNAPTIC VESICLES).
GN KIF1A OR ATSV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96299637; PubMed=8661001;
RA Furlong R.A., Zhou C.Y., Ferguson-Smith M.A., Affara N.A.;
RT *Characterization of a kinesin-related gene ATSV, within the tuberous
sclerosis locus (TSC1) candidate region on chromosome 9q34.*;
RL Genomics 33:421-429(1996).
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC
VESICLE PRECURSORS (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: X90840; CAA62346.1; -.
CC HSSP: P17119; 3KAR.
CC MIM: 601255; -
DR InterPro; IPR000253; -
DR InterPro; IPR001752; -
DR InterPro; IPR001849; -
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil.
KW DOMAIN 1 361 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 366 383 COILED COIL (POTENTIAL).
FT DOMAIN 429 462 COILED COIL (POTENTIAL).
FT DOMAIN 516 572 FHA.

ALIGNMENTS

```
FT DOMAIN 622 681 COILED COIL (POTENTIAL).
FT DOMAIN 801 822 COILED COIL (POTENTIAL).
FT DOMAIN 1575 1673 PH.
FT NP_BIND 97 104 ATP (POTENTIAL).
SQ SEQUENCE 1690 AA; 191083 MW; D8DDEC784624FB4D CRC64;

Query Match 62.8%; Score 1152; DB 1; Length 1690;
Best Local Similarity 62.1%; Pred. No. 1.3e-75;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;

Qy 4 CGNKKVVVVRPFNAREIDRGAKCIVRMENQNTILTPPPGAEKARKSGKTIMDGPKAPA 63
Db 3 GASYKVVAVRPFNSREMSRDSKCIQMSGSTTVVNPQPKET-----PKSFS 51

Qy 64 FDRSYWSPDKNAP---NYARQEDLFQDLGVPLDNAPFGYNNCFAYGQTGSGKYSMMG 120
Db 52 FDYSYWS--HTSPEDINVASOKQYVRDGEMLQAFEGYNVCIFAYGQTGAGKSYTMWG 109

Qy 121 YGK--EHGVIPRICODMFRINELQKDNLTCTVEVSYLEIYNERNVRLDLPSTKGNLKV 178
Db 110 KQEKDQOQGIIPQLCEDLFSRINDTND-NMSYSVEVSMEIYCVSRVRLDLPKKNGLRV 168

Qy 179 REHPSTGPIVEDLAKLVVRSQETENLDEGNKARTVAATNMNETSSRSASHVFTLTQK 238
Db 169 REHPLGPIYVEDLSKLVTSYNDIQDLDSGNKPRTPVAATNMNETSSRSASHVFTLTQK 228

Qy 239 WHDETNDKTEKVAKISLVLDLAGSERATSTGATGARLKEGAEINRSLSLTGLGRVIAALADM 298
Db 229 RHDATNTTTEKVKISLVLDLAGSERADSTGAKTRLEKAGININKSLTTLGKVISALAE 288

Qy 299 SSG-----KQKNQIVPYRDSVLTLWLLKDSLGNSMTAMIAISPADINFEETLSTLYA 353
Db 289 DSGPNKNNKKKTDPIPYRDSVLTLWLLRENLGNSRTAMVAALSPADINFEETLSTLYA 348

Qy 354 DSAK 357
Db 349 DRAK 352

RESULT 2
KF1A_MOUSE STANDARD; PRT; 1695 AA.
AC P3173; O61770;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1A.
GN KIF1A OR KIF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=95292344; PubMed=7539720;
RA Okada Y., Yamazaki H., Sekine-Aizawa Y., Hirokawa N.;
RT "The neuron-specific kinesin superfamily protein KIF1A is a unique
RT monomeric motor for anterograde axonal transport of synaptic vesicle
RT precursors."
RL Cell 81:769-780(1995).
RN [2]
RP PRELIMINARY SEQUENCE OF 100-247 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=93077686; PubMed=1447303;
RA Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RA Hirokawa N.;
RT "Kinesin family in murine central nervous system."
RL J. Cell Biol. 119:1287-1296(1992).
CC -1- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC
CC VESICLE PRECURSORS.
CC -1- SUBUNIT: MONOMER.
```


RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AX039604; CAC16629.1; -;
DR EMBL; AB011163; BAA25517.1; -;
DR EMBL; AK029777; BAB14341.1; -;
DR InterPro; IPR000253; -;
DR InterPro; IPR001849; -;
DR Pfam; PF00169; PH; 1;
DR Pfam; PF00498; FHA; 1;
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1;
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1;
DR PROSITE; PS00006; FHA_DOMAIN; 1;
DR PROSITE; PS50003; PH_DOMAIN; 1;
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 361
FT DOMAIN 365 386
FT DOMAIN 470 502
FT DOMAIN 556 612
FT DOMAIN 668 737
FT DOMAIN 841 869
FT DOMAIN 1702 1799
FT NP_BIND 97 104
SQ SEQUENCE 1816 AA; 204474 MW; AD62F0515978C783 CRC64;

Query Match 61.08; Score 1119; DB 1; Length 1816;
Best Local Similarity 60.88; Pred. No. 3.7e-73;
Matches 220; Conservative 59; Mismatches 63; Indels 20; Gaps 5;
5;
QY 4 GGNIKVVVRVPPNAREIDGAKCIVRMENQOTILTPPGAEERKARKSGKTMDGPKAFA 63
DB 3 GASVKVAVRVPFNSRETSKESKILIQMGNSTSIINPKNPK-----APKSFS 51
QY 64 FDRSYWSF-DKNAPNYARQEDLFDGLVPLLDNAFKYNNCIFAYGTGSGKSYSMAGY 122
DB 52 FDYSYWSHTSPEDCFASQNRVYNDIGKMLLHAFEGYVNCIFYAGTGAGKSYTMGKQ 111
QY 123 KEH--GVIPRICQDMFRINELQDKNLCTVEYSYLEIYNERYVRLDNLSTKGNLVRE 180
DB 112 EESAGIIPQCEEFKEKIND-NCNEEMSYSVESYMEYICERVRLDNLNPKGNLVRE 170
QY 181 HPSTGPYVEDLAKLVVRSFQBIENLDEGNKARTVAATNNMETSRSRSHAVFTLTQKWH 240
DB 171 HPLGPGYVEDLSKLVASYTYTDIADLMDAGNKARTVAATNNMETSRSRSHAVFTIYTKKH 230
QY 241 DEETKMDTEKVAKLSLVDLAGSERATSTGATGARKLKEGAEINRSLSTLGRVIAALDM-- 298
DB 231 DNETNLSTKYSKLSLVDLAGSERADSTGAKTRLKEGANINKSLITLIGKVISALAFVDN 290
QY 299 ---SSGOKKNQLPYRDSVLTWLLKSLGNSMTAMAAISPADINFEETLSLRLVADS 355
DB 291 CTSKSKKKKTKDFIPYRDSVLTWLLRENLGNSRTAMVAALSPADINYDETLSLRYADR 350
QY 356 AK 357

Db 351 AK 352
II
RESULT 5
KF1B_MOUSE
ID KF1B_MOUSE STANDARD; PRT; 1816 AA.
AC Q60575; Q9WE5; Q9R0B4; Q9Z119;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Brain;
RX MEDLINE=95094296; PubMed=7528108;
RA Nangaku M., Sato-Yoshitake R., Okada Y., Noda Y., Takemura R.,
RA Yamazaki H., Hirokawa N.;
RT "KIF1B, a novel microtubule plus end-directed monomeric motor protein
RL for transport of mitochondria.";
RL Cell 79:1209-1220(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=ICR; TISSUE=Brain;
RX PubMed=10571041;
RA Gong T.L., Winnicki R.S., Kohrman D.C., Lomax M.I.;
RT "A novel kinesin of the UNC-104/KIF1 subfamily encoded by the Kif1b
RT gene.";
RL Gene 239:117-127(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,
RA Perry V.H., Coleman M.P.;
RT "The major brain isoform of Kif1b lacks the putative mitochondria-
RL binding domain.";
RL Mamm. Genome 10:0-0(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=ICR;
RA Nakagawa T., Hirokawa N.;
RT "Identification and characterization of a new kinesin superfamily
RT KIF1B-beta.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC MICROTUBULE PLUS END-DIRECTED MOTILITY.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2/BETA AND 3;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
CC TISSUE (MAINLY IN THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE
CC TYPE OF NEURONAL CELL.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17577; BAA04503.1; -;


```

FT DOMAIN 777 797 COILED COIL (POTENTIAL).
FT DOMAIN 183 335 PH.
FT DOMAIN 1460 1558 PH.
FT DOMAIN 957 1052 ARG/LYS-RICH (BASIC).
FT DOMAIN 1203 1584 ARG/LYS-RICH (BASIC).
FT NP_BIND 93 100 ATP (BY SIMILARITY).
FT VARIANT 598 598 I -> T.
FT VARIANT 930 930 V -> M.
SQ SEQUENCE 1584 AA; 179736 MW; 559462B3FD029B43 CRC64;

Query Match 60.3%; Score 1106.5; DB 1; Length 1584;
Best Local Similarity 61.2%; Pred. No. 2.4e-72;
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

QY 6 NIKVVVRPPNAREIDRGAKCIVRMGNQTLTPPPGAEEKARKSGKTMIDGPKAFD 65
DB 3 SVKAVVRPPNAREISNTSKVLQVNGTTT-----NGHSINKENFSNFD 50

QY 66 RSYWSEFNAPNARQEDLFQDLGVPLLDNAFKYNNCIFYAGTGGSGKSYMMGYGK-- 123
DB 51 HSYWSEFNAPNARQEDLFQDLGVPLLDNAFKYNNCIFYAGTGGSGKSYMMGKANDP 110

QY 124 -EHGVIPRICQDMFRINELQKKNLTCTVEVSYLEIYNERVDRLLNPFSTGNLKVREHP 182
DB 111 DEMGIIPRLCNDLFARDN--NNDKDVQVSEVSYMEIYCEYKRVKLLNPNSGNLRVREHP 169

QY 183 STGPYVEDLAKLVRSFOEINLMDGKARTVAATNMNSTRSSSHAVFTLTOKWHDE 242
DB 170 LLGYPVDDLTKMAYCSYHDICNLMDGKARTVAATNMNSTRSSSHAVFTLTOKWHCA 229

QY 243 ETKMDTEKVAKISLVLDLAGSERATSTGATGARLKEGAEINSLTGLRVIAALADMSSGK 302
DB 230 DSNLDEKHSKISLVLDLAGSERATSTGATGARLKEGAEINSLTGLRVIAALADMSSGK 289

QY 303 QKKNQ-LVPYRDSVLTWLLKDLGSGNSMTAMIAAISPADINFEETLSLTRYADSAK 357
DB 290 KKSNGVPIYRDSVLTWLLKDLGSGNSMTAMIAAISPADINFEETLSLTRYADRAK 345

RESULT 7
KF1B_RAT ID KF1B_RAT STANDARD; PRT; 1097 AA.
AC O35787;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1D.
GN KIF1D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC STRAIN=WTSTAR;
RA Rogers K.R., Griffin M., Brophy P.J.;
RT "The secretory epithelial cells of the choroid plexus employ a novel
kinesin-related protein.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ000696; CA04248.1; -.

```

```

DR HSSP; P17119; 3KAR.
DR InterPro; IPR000253; -.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; kinesin; 1.
DR Pfam; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS00006; FHA_DOMAIN; FALSE_NEG.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 357 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 358 380 COILED COIL (POTENTIAL).
FT DOMAIN 437 478 COILED COIL (POTENTIAL).
FT DOMAIN 520 587 FHA.
FT DOMAIN 630 671 COILED COIL (POTENTIAL).
FT DOMAIN 824 868 COILED COIL (POTENTIAL).
FT NP_BIND 96 103 ATP (POTENTIAL).
SQ SEQUENCE 1097 AA; 122333 MW; 8EF40B1C7579BA5B CRC64;

Query Match 59.2%; Score 1085.5; DB 1; Length 1097;
Best Local Similarity 60.5%; Pred. No. 5e-71;
Matches 216; Conservative 64; Mismatches 60; Indels 17; Gaps 7;

QY 4 GGNIKVVVRPPNAREIDRGAKCIVRMGNQTLTPPPGAEEKARKSGKTMIDGPKAFA 63
DB 3 GASVYAVVRPPNARETSQDAKCVVSMQNTTSIINP-----KQSRMFL---KA-S 50

QY 64 FDRSYWSEFDK-NAPNARYAQEDLFQDLGVPLLDNAFKYNNCIFYAGTGGSGKSYMMGYG 122
DB 51 FDYSWHSVTSVEDPQFASQQVYRDIGEEMLLHAFEGYVNCIFYAGTGGSGKSYMMGRQ 110

QY 123 K--EHGVIPRICQDMFRINELQKKNLTCTVEVSYLEIYNERVDRLLNPFSTGNLKVRE 180
DB 111 EPGQQGVPQPCEDLFNRV--VNSQAQLSYSEVSYMEIYCEYKRVKLLNPNKSRGSLRYRE 169

QY 181 HPSTGPYVEDLAKLVRSFOEINLMDGKARTVAATNMNSTRSSSHAVFTLTOKWH 240
DB 170 HPILGPYVDLKLAVTSYADIADLMDGKARTVAATNMNSTRSSSHAVFTLTOKWH 229

QY 241 DEETKMDTEKVAKISLVLDLAGSERATSTGATGARLKEGAEINSLTGLRVIAALADMSS 300
DB 230 DQLTGLDSEKYSKISLVNLAGSERADSSGARGMRLKEGANINKSLTTLGKVISALADLQS 289

QY 301 GKQKNQLVYRDSVLTWLLKDLGSGNSMTAMIAAISPADINFEETLSLTRYADSAK 357
DB 290 -KRRKSDFIYRDSVLTWLLKDLGSGNSRTAMIAAISPADINFEETLSLTRYADRTK 345

RESULT 8
KF1B_RAT ID KF1B_RAT STANDARD; PRT; 689 AA.
AC O88658;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B (FRAGMENT).
GN KIF1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RA Faire K., Gruber D., Bulinski J.C.;
RT "Identification of kinesin-like molecules in myogenic cells.";
RL Eur. J. Cell Biol. 0:0-0(1998).
CC -1- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
SUBFAMILY.
CC -----

```



```
CC CC -----
CC CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AF083331; AAC33292.1; -.
CC CC HSSP; P33176; I8G2.
CC CC InterPro; IPR000253; -.
CC CC InterPro; IPR001356; -.
CC CC InterPro; IPR001752; -.
CC CC Pfam; PF00225; kinesin; 1.
CC CC Pfam; PF00498; FHA; 1.
CC CC PRINTS; PR00380; KINESINHEAVY.
CC CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
CC CC PROSITE; PS0067; FHA_DOMAIN; FALSE_NEG.
CC CC PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
CC CC Motor protein; Microtubules; ATP-binding; Coiled coil.
CC CC DOMAIN 1 359 MECHANOCHEMICAL (MOTOR).
CC CC FT DOMAIN 360 381 COILED COIL (POTENTIAL).
CC CC FT DOMAIN 451 483 COILED COIL (POTENTIAL).
CC CC FT DOMAIN 538 618 FHA.
CC CC FT DOMAIN 651 688 COILED COIL (POTENTIAL).
CC CC FT NP_BIND 97 104 ATP (POTENTIAL).
CC CC FT NON_TER 689
CC CC SEQUENCE 689 AA; 77352 MW; F64818C12FE69319 CRC64;
CC CC -----
CC CC Query Match 58.5%; Score 1073.5; DB 1; Length 689;
CC CC Best Local Similarity 60.2%; Pred. No. 2e-70;
CC CC Matches 216; Conservative 57; Mismatches 67; Indels 19; Gaps 7;
CC CC -----
CC CC 4 GGNKVVVVRFPNAREIDRGAKIVRM-EGNQITLTTPPGAEKARKSGKTMDGPKA 63
CC CC 3 GASVAVVRFPNAREIDRGAKIVRM-EGNQITLTTPPGAEKARKSGKTMDGPKA 63
CC CC -----
CC CC 64 FDRSYWSP-DKNAPNARYQEDLFDQGLVPLDNLAFKGNKNCIFAYGQTGSGKSYSMXG 122
CC CC 52 FDYSYNSHTSPDCPQASQSRVYNDIGKEMLLHAFEGYNYCIVFAYGQTGAGKSYTMGKQ 111
CC CC -----
CC CC 123 KEH--GVIPRICODMFRINELQKDKNLTCTVEVSYLEYINRVRDLNPNSTKGNLKVSE 180
CC CC 112 EESQAGIITPTCEELFEKIND-NCNEDMSYVS-SYMEITCEKVRDLNPNKGNLKVRE 169
CC CC -----
CC CC 181 HPSTGPYVEDLAKLVRSFOEINLMDGNKARTVAATNMNETSSRSRSHAVFTLTQKWH 240
CC CC 170 HPLGPGYVEDLSKLVTSYTDIALDLADGAKNARTVAATNMNETSSRSRSHAVFTVTKQK 229
CC CC -----
CC CC 241 DEETKMDTEKVAKISLVDLAGSRATSTGATGARLKEGAEINRSLSLGRVIALADMS 300
CC CC 230 DPETNLSTKTVKISLVDLAGSRADSTGATGARLKEGRNINKSITTLCKVISLAEVSK 289
CC CC -----
CC CC 301 GKQKKNQVYRDSVLTWMLKDSLGNSMTAMIAISPADINFEETL--STLRVADSAK 357
CC CC 290 -KKKKTDFIPYRDSVLTWMLRENGLGNSRTAMVAALSPADINYEETLSTLRVADRAK 347
CC CC -----
CC CC RESULT 9
CC CC K122_STRPU STANDARD; PRT; 699 AA.
CC CC AC P46872;
CC CC DT 01-NOV-1995 (Rel. 32, Created)
CC CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC CC DE KINESIN-II 85 KDA SUBUNIT (KRP-85/95 85 KDA SUBUNIT).
CC CC GN KRP85.
CC CC OS Strongylocentrotus purpuratus (Purple sea urchin).
CC CC OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
CC CC Echinoidae; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
```

```
OC Strongylocentrotus.
OX NCBI_taxid=7688;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE-Egg;
RX MEDLINE=94050179; PubMed=8232586;
RA Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,
RA Scholey J.M.;
RA "Novel heterotrimeric kinesin-related protein purified from sea
RA urchin eggs.";
RL Nature 366:268-270(1993).
CC CC -!- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
CC CC OF 95 AND 85 KDA.
CC CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC CC II SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; L16993; AAA16098.1; -.
CC CC HSSP; P17119; 3KAR.
CC CC InterPro; IPR001752; -.
CC CC Pfam; PF00225; kinesin; 1.
CC CC PRINTS; PR00380; KINESINHEAVY.
CC CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC CC PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
CC CC Motor protein; Microtubules; ATP-binding; Coiled coil.
CC CC DOMAIN 1 340 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC CC FT DOMAIN 341 619 COILED COIL (BY SIMILARITY).
CC CC FT DOMAIN 620 699 GLOBULAR (BY SIMILARITY).
CC CC FT NP_BIND 97 104 ATP (POTENTIAL).
CC CC SEQUENCE 699 AA; 78697 MW; 7B3866111CB08190 CRC64;
CC CC -----
CC CC Query Match 40.5%; Score 743.5; DB 1; Length 699;
CC CC Best Local Similarity 46.0%; Pred. No. 1.7e-46;
CC CC Matches 167; Conservative 62; Mismatches 99; Indels 35; Gaps 10;
CC CC -----
CC CC 2 SGGGNKVVVVRFPNAREIDRGAKIVRM-EGNQITLTTPPGAEKARKSGKTMDGPK 60
CC CC 6 SGNDNRVVVVRCPPLSKSETGGTGGFKSVVKMDMRGTGVQTPNPA-----PSG----EPPK 56
CC CC -----
CC CC 61 AFAPDRSYWSPDKNAPNARYQEDLFDQGLVPLDNLAFKGNKNCIFAYGQTGSGKSYSMXG 120
CC CC 57 SFTFDVTP-----APG-AKQTDVYNQTPARPIDVAIEGYNGTIFAYGQTGKTFTMEG 109
CC CC -----
CC CC 121 YCKE---HGVIPRICODMFRINELQKDKNLTCTVEVSYLEYINRVRDLNPNSTKGNL 177
CC CC 110 VRSQDELGLIIPNAPFHGHI--AKQENVRFLVRSVLEYINRVRDLNPNSTKGNL 167
CC CC -----
CC CC 178 VREHSTGPYVEDLAKLVRSFOEINLMDGNKARTVAATNMNETSSRSRSHAVFTLTQ 237
CC CC 168 VKERPDGVVYVKDLSAFVYNNADDMDRITLGNKNSVYGATNMNNESSRSHALFTITLER 227
CC CC -----
CC CC 238 KWRDEETKMDTE---KVAKISLVDLAGSRATSTGATGARLKEGAEINRSLSLGRVIA 294
CC CC 228 ----SDMGLDKEQHVVRGKLVHVDLAGSRQTVGTATGQRLKEATKINLSLTGNVSS 283
CC CC -----
CC CC 295 LADSSGKQKNQVYRDSVLTWMLKDSLGNSMTAMIAISPADINFEETLSTLRVAD 354
CC CC 284 LVD-----GKSTHPIYRNSKLTLLQDLSLGNKNTVMCANIGPAEYNYDETISLRYAN 337
CC CC -----
CC CC 355 SAK 357
CC CC 338 RAK 340
```

```
RESULT 10
K121_STRPU STANDARD; PRT; 742 AA.
AC P46871;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE KINESIN-II 95 KDA SUBUNIT (KRP-85/95 95 KDA SUBUNIT).
GN KRP95.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=94050179; PubMed=8232586;
RA Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,
RA Scholey J.M.;
RT "Novel heterotrimeric kinesin-related protein purified from sea
RT urchin eggs";
RL Nature 366:268-270(1993).
CC -1- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
CC OF 95 AND 85 KDA.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U09596; AAA87393.1; -
CC HSSP; P56536; 2KIN.
CC InterPro: IPR001752; -.
CC Pfam; PF00225; kinesin; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil.
KW DOMAIN 1 337 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 338 613 COILED COIL (BY SIMILARITY).
FT DOMAIN 614 742 GLOBULAR (BY SIMILARITY).
FT NP_BIND 95 102 ATP (POTENTIAL).
SQ SEQUENCE 742 AA; 84202 MW; 47C40A367BAA7B5 CRC64;

Query Match 38.8%; Score 712; DB 1; Length 742;
Best Local Similarity 45.3%; Pred. No. 3.5e-44;
Matches 163; Conservative 56; Mismatches 101; Indels 40; Gaps 9;

QY 7 IKVVVRPPFNAREIDRCAGKIVRMENQTI--LTTPPGAEKARKSGKTTMDGPKAFAP 64
DB 9 VKVVRCRPMNSKSEIKGRKRVEMDNKRGLVEVTPKGPPEPNKS-----FTF 58
QY 65 DRSY-WSPDKNAPNVAQEDLFQDLGVPLLDNAFGVNNCIFAYGQTGSGKSYSMVGK 123
DB 59 DTVDWN-----SKQIDLYDETFSLVESVLQGFNGTIFAYGQTGKTFTMEGVR 110
QY 124 E---HGVIPTCDMFRINELQKDKNLTCTVEVSYLEIYNNVRVDLNPSTKGNLKVRE 180
DB 111 NPELRGVIPNSFEHIFTHIARTNQOFL---VRASYLEIYQEEIRDLAKDQKKRLDLKE 167
QY 181 HPSGTPYVEDLAKLVRSFQBIENLMEGNKARTVAATNMNETSRSRHAVFTLITQKWH 240
DB 168 RPTGVIYKDLSSVFTSVKSEIHEVMTVNNNRVSGVNNMNEHSHAIPIITI----- 223
QY 241 DEETKMDTE---KVAKISLVDLGASRATSTGATGARLKEAGAEINRSLSLTLGRVIAALAD 297
```

```
Db 224 CSELGVDSGENHVRGKLNVLDSAGSERQAKTGATGDRLEKATKLNLSALGNVISALVD 283
QY 298 MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFTLTSLTRYADSAK 357
DB 284 -----GKSSHIPYRDSKLTLLQDSLGGNAKTVMVANMGPASYNFDETTITTRYANRAK 337

RESULT 11
KINL_LEICH STANDARD; PRT; 955 AA.
AC P46865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE KINESIN-LIKE PROTEIN K39 (FRAGMENT).
GN KIN.
OS Leishmania chagasi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5669;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/BR/82 / ISOLATE BA-2;
RX MEDLINE=93133867; PubMed=8421715;
RA Burns J.M. Jr., Shreffler W.G., Benson D.R., Chalib H.W., Badaro R.,
RA Reed S.G.;
RT "Molecular characterization of a kinesin-related antigen of
RT Leishmania chagasi that detects specific antibody in African and
RT American visceral leishmaniasis.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
CC -1- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L07879; AAA29254.1; -
CC HSSP; P17119; 3KAR.
CC InterPro: IPR001752; -.
CC Pfam; PF00225; kinesin; 1.
CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil; Repeat.
KW DOMAIN 1 399 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 426 >955 COILED COIL (POTENTIAL).
FT NP_BIND 122 129 ATP (POTENTIAL).
FT DOMAIN 704 >955 7 X 39 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 704 742 1.
FT REPEAT 743 781 2.
FT REPEAT 782 820 3.
FT REPEAT 821 859 4.
FT REPEAT 860 898 5.
FT REPEAT 899 937 6.
FT REPEAT 938 >955 7 (PARTIAL).
FT NON_TER 955 955
SQ SEQUENCE 955 AA; 106168 MW; 8CA76815BBE84C6E9 CRC64;

Query Match 38.5%; Score 705.5; DB 1; Length 955;
Best Local Similarity 42.7%; Pred. No. 1.4e-43;
Matches 163; Conservative 58; Mismatches 126; Indels 35; Gaps 8;

QY 7 IKVVVRPPFNARE--IDRGAKIVRMENQTI--LTTPP-----PGAEKARKSGKTTMD 57
DB 13 VKVSVVRPLNERENNAPEGTKVTVAQAAAVTVKVLGGSNNSGAAESMGTAARRVAQD 72
QY 58 GPKAFADRSYWSF---DKNAPNTVAQEDLFQDLGVPLLDNAFGVNNCIFAYGQTGSGK 114
```

Db 73 ----FOFDIVFVSVETPDAGCATPATQDVFRTIGVPLVQIHAEDGFNSCLFAYGQTSGK 128
 Qy 115 SYSMG-----YKHEGVIPRICODMFRINELQKDKNLCTCTVEVSLIYINERVDLL 168
 Db 129 TYTMMGADYSALSGEGNGVTPRICLETFARKASVEAQGHSEWIVELGYVYVNERVSDLL 188
 Qy 169 NPSTKG-----NLKVRHPSGTGPVVDLAKLVVRSFQEIENLMDEGNKARTVAATNM 220
 Db 189 GKRRKGVKGGEVYVDVREHPSRGVFLGORLVEVGSLLDDVRLIEICNGVRHTASTKM 248
 Qy 221 NEYSSRSHAVFTLTL-----TQKWHDEETKMDTEKAKISLVDLAGSERATSTGATGARLK 276
 Db 249 NDNSKSHALIMLLREERTMTTKSETITTAGKSRMNLVDLAGSERVAQSQVEGGQFK 308
 Qy 277 EGAEINRSLSLTGRVTAALDNSS--GKOKNLQVLPYRDSVLTWLLKDSLGGNSMTAMIAA 335
 Db 309 EATHINLSLTTIGRVIDVADLMATKGAQAYSVAPFRDSKLTFLKDSLGGNSKTFMIAT 368
 Qy 336 ISPADINFEETLSRLYADSAK 357
 Db 369 VSPSALNYEETLSRLVRSAR 390

 RESULT 12
 KIF3B_HUMAN
 ID KIF3B_HUMAN STANDARD; PRT; 747 AA.
 AC O15066;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN MOTOR 3B) (HH0048).
 GN KIF3B OR KIAA0359.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxId-9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RX MEDLINE-97349984; PubMed-9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
 RT DNA Res. 4:141-150(1997).
 RL -1- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING ACTIVITY IN VITRO (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN II SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB002357; BAA20815.1; -
 CC HSSP; P56536; 2KIN.
 CC MIN; 603754; -
 CC InterPro; IPR001752; -
 CC Pfam; PF00225; kinesin.1.
 CC PRINTS; PR00380; KINESINHEAVY.
 CC PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 CC Motor protein; Microtubules; ATP-binding; Coiled coil; Neurene.
 KW DOMAIN 1 345 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
 FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
 FT DOMAIN 580 747 GLOBULAR.


```
Db 3 GASVKVAVRVPFPNARETSQDAKCVVSMQNTTISINP-----KQSKDAPKSF 51
Qy 64 FDRSYWSPDKNA-DNYARQEDLFDGLVPLLDNAFKYNNCIFAYGQSGKSYMMGYG 122
Db 52 FDSYWSHTSTEDPOFASQQQVYRDIGEMLLHAFEGYNNVCIFAYGQSGKSYMMGRQ 111
Qy 123 K--EHGVIPRICODMFRINELQDKNLTCTVEVSYLEIYNERNVRLDLPSTKGNLKVRE 180
Db 112 EPQGGIVPOLCEDLFSRVSENO-SAQLSYSVEVSMEIYCERYVDLLNPKSRGSLVRE 170
Qy 181 HPSTGYPVEDIAKLIVRSFQFIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKWH 240
Db 171 HPILGYPVQDLASKLAVTSYADIADLMDCGNKARTVAATNMNETSSRSRHAVFTIVTQRCH 230
Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300
Db 231 DQTLGLDSEKYSKISLVDLAGSERADSSGARGMKLKEGANINKSLTTLGKVISALADMS 290
Qy 301 GKQKNOLPVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADSAK 357
Db 291 -KKKSTDFIPYRDSVLTWLLKENLGGNSRTAMIAALSPADINYEETLSTLRYADRTK 346

RESULT 2
Q92119 ID Q92119 PRELIMINARY; PRT; 1770 AA.
AC Q92119
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KIF1B-BETA.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR MOUSE;
RA Nakagawa T., Hirokawa N.;
RT "Identification and characterization of a new kinesin superfamily
KIF1B-beta.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023656; BAA75243.1; -.
DR HSP; P33176; IBC2.
DR INTERPRO; IPR000255; -.
DR INTERPRO; IPR001752; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1770 AA; 198850 MW; E444228501672E19 CRC64;
```

Query Match 60.7%; Score 1114; DB 11; Length 1770;
Best Local Similarity 61.3%; Pred. No. 2.5e-73;
Matches 219; Conservative 59; Mismatches 63; Indels 16; Gaps 5;

```
Qy 4 GGNKVVVVRVPFPNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIIMDGPKAFA 63
Db 3 GASVKVAVRVPFPNSRSTKESKCIIOGNSSTIINPKNPK-----APKSFS 51
Qy 64 FDRSYWSPDKNA-PNARYARQEDLFDGLVPLLDNAFKYNNCIFAYGQSGKSYMMGYG 122
Db 52 FDSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNNVCIFAYGQSGKSYMMGK 111
Qy 123 KEH--GVIPRICODMFRINELQDKNLTCTVEVSYLEIYNERNVRLDLPSTKGNLKVRE 180
```

```
Db 112 EESQAGIIPOLCEELFEKIND-NCNEEMSYSVEVSMEIYCERYVDLLNPKNKNLRYRE 170
Qy 181 HPSTGYPVEDIAKLIVRSFQFIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKWH 240
Db 171 HPLGLPYVEDLSKLAVTSYTDIADLMDCGNKARTVAATNMNETSSRSRHAVFTIVTQKKQ 230
Qy 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAENRSLSTLGRVIAALADMS 300
Db 231 DPETNLSTKESKYSKISLVDLAGSERADSTGAKTRLKEGANINKSLTTLGKVISALAEYSK 290
Qy 301 GKQKNOLPVPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADSAK 357
Db 291 -KKKSTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYEETLSTLRYADRAK 346

RESULT 3
Q9ROB4 ID Q9ROB4 PRELIMINARY; PRT; 1816 AA.
AC Q9ROB4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR OUTBRED; TISSUE=BRAIN;
RA Gong T.L., Winnicki R.S., Kohrman D.C., Lomax M.I.;
RT "A novel kinesin of the UNC-104/KIF1 subfamily encoded by the kif1b
gene.";
RL Gene 239.117-127 (1999).
DR EMBL; AF090190; AAF06718.1; -.
DR HSP; P33176; IBC2.
DR MGD; MGI:108426; Kif1b.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1816 AA; 204079 MW; E316EC295138E5DE CRC64;
```

Query Match 60.5%; Score 1110; DB 11; Length 1816;
Best Local Similarity 60.5%; Pred. No. 5.1e-73;
Matches 219; Conservative 59; Mismatches 64; Indels 20; Gaps 5;

```
Qy 4 GGNKVVVVRVPFPNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIIMDGPKAFA 63
Db 3 GASVKVAVRVPFPNSRSTKESKCIIOGNSSTIINPKNPK-----APKSFS 51
Qy 64 FDRSYWSPDKNA-PNARYARQEDLFDGLVPLLDNAFKYNNCIFAYGQSGKSYMMGYG 122
Db 52 FDSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNNVCIFAYGQSGKSYMMGK 111
Qy 123 KEH--GVIPRICODMFRINELQDKNLTCTVEVSYLEIYNERNVRLDLPSTKGNLKVRE 180
Db 112 EESQAGIIPOLCEELFEKIND-NCNEEMSYSVEVSMEIYCERYVDLLNPKNKNLRYRE 170
Qy 181 HPSTGYPVEDIAKLIVRSFQFIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKWH 240
Db 171 HPLGLPYVEDLSKLAVTSYTDIADLMDCGNKARTVAATNMNETSSRSRHAVFTIVTQKKQ 230
```


OS *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodoridae; Pelodirinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Therry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*
RT *elegans*.;
RT Nature 368:32-38(1994).
RX [2]
RP SEQUENCE FROM N.A.
RX Du Z.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RA	Waterston R.;	SEQUENCE FROM N.R.
DR	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; U50135; AAA93453.1; -,	
DR	HSSP; P17119; 3KAR.	
DR	INTERPRO; IPR001752; -,	
DR	INTERPRO; IPR001849; -,	
DR	PFAM; PF001169; PH: 1.	
DR	PFAM; PF002225; kinesin: 1.	
DR	PRINTS; PR00380; KINESINHEAVY.	
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.	
DR	PROSITE; PS00003; PL_DOMAIN; 1.	
DR	PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.	
DR	Motor protein; Microtubules; ATP-binding; Colled coll.	
SW	Motor protein; Microtubules; ATP-binding; Colled coll.	
KW	SEQUENCE 1584 AA; 179649 NW; C06DBI76AC9C8454 CRC64;	

[illegible]

	Db	230	DSNLDTEKHSKISLVDLAGSERANSTGAGQRLKEGANINKSLTTLGLVLSKLAEESTKK	289
	QY	303	QKKNQ-LVPYRSVLTWLLKDSLGGNSMTAMTAISPADINFEETLSTLYRADSAK	357
			K : N : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :	
	Db	290	KKSNGKVIPIYRDSVLTWLLLRNLGGNSKTAMLAAISPADINFEDETLSLTLYADRAK	345
	RESULT	6		
	Q9V7TG			

ID Q9V7T6 PRELIMINARY; PRT; 1773 AA.
 AC Q9V7T6;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE CG8566 PROTEIN.
 GN CG8566.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003805; AAF57957.1; -.
 DR HSP; P17119; 3KAR
 DR FLYBASE; FBgn0034155; CG8566.
 DR INTERPRO; IPR000253; -.
 DR INTERPRO; IPR001752; -.
 DR INTERPRO; IPR001849; -.
 DR PFAM; PF00169; PH; 1.
 DR PFAM; PF00225; kinesin; 1.
 DR PFAM; PF00498; FHA; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 SQ SEQUENCE 1773 AA; 200755 MW; 335BESCD5E4DB681 CRC64;

Query Match 60.3%; Score 1106.5; DB 5; Length 1773;
 Best Local Similarity 62.4%; Pred. No. 8.9e-73;
 Matches 222; Conservative 52; Mismatches 69; Indels 13; Gaps 4;

QY 6 NIKVVVRVPENAREIDRGAKCIIVRMGNQITLPPPGAEEKARKSGKTIIMDGKAFAD 65
 DB 32 SVKVAVRVPENAREIDRGAKCIIVRMGNQITLPPPGAEEKARKSGKTIIMDGKAFAD 83
 QY 66 RSYWSFDKNAFYARQEDLFQDLGVPLIDNFKYNNICFYAGTGTSGKSYSMGKRGK- 124
 DB 84 YSYWSHDHDDADFTQSMVYKNDIGCEMLQHSFDGYNVCIFYAGTGTGKSYSMGKRGK 143
 QY 125 -HGVPTRICQDMFRINELQDKNLTCTVVSYLEIYNNRVDLLNPSKGNLKVREHPS 183
 DB 144 QEGTIPMICKDLFTRIQTETD-DLKYSVEVSMEIYCEYRVDLLNPKNKNLVRHPL 202
 QY 184 TGPVVDLAKLVRSFQEIENLMDGNKARTVAATNNMETSRSRSHAVETLTLTKQWDEE 243
 DB 203 LGPYVEDLSKLVADTDYQDIDHLDGNKARTVAATNNMETSRSRSHAVETLTLTKQWDEE 262
 QY 244 TKMTEKVAKISLVDLAGSERATSTGATGARKLKGAEINRSLTSLGRVTAALADMSKQ 303
 DB 263 TNLITKRVKISLVDLAGSERADSTGATGARKLKGAEINRSLTSLGRVTAALADMSKQ 322
 QY 304 --KKNQLVPYRDSVLTWLLKSLGNSMTAMIAISPADINFEETLSFLRYADSAK 357
 DB 323 NTKKADFPYRDSALTWLLRENLGNSKMTAMIAISPADINFEETLSFLRYADSAK 378
 RESULT 7
 Q9WVE5 PRELIMINARY; PRT; 1816 AA.
 ID Q9WVE5
 AC Q9WVE5;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE KIF1B MAJOR ISOFORM.
 GN KIF1B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=BRAIN;
 RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,
 RA Perry V.H., Coleman M.P.;
 RT "The major brain isoform of Kif1b lacks the putative mitochondria-
 RT binding domain.";
 RL Mamm. Genome 10:0-0(1999).
 DR EMBL; AF131865; AAD39438.1; -.
 DR HSP; P33176; 18C2
 DR MGD; MGI:108426; Kif1b.
 DR INTERPRO; IPR000253; -.
 DR INTERPRO; IPR001752; -.
 DR INTERPRO; IPR001849; -.
 DR PFAM; PF00169; PH; 1.
 DR PFAM; PF00225; kinesin; 1.
 DR PFAM; PF00498; FHA; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.
 SQ SEQUENCE 1816 AA; 204153 MW; 95CB196A9DB4895A CRC64;

Query Match 60.1%; Score 1102; DB 11; Length 1816;
 Best Local Similarity 60.2%; Pred. No. 2e-72;
 Matches 218; Conservative 59; Mismatches 65; Indels 20; Gaps 5;

QY 4 GGNKVVVRVPENAREIDRGAKCIIVRMGNQITLPPPGAEEKARKSGKTIIMDGKAFAD 63
 DB 3 GASVKVAVRVPENAREIDRGAKCIIVRMGNQITLPPPGAEEKARKSGKTIIMDGKAFAD 83

[illegible]

```

RN      [1]  ...AD-10126;
RP      SEQUENCE FROM N.A.
RC      STRAIN=WLSTAR;
RL      Rogers K.R., Griffin M., Brophy P.J.;
RM      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RS      EMBL: AJ000696; CAA04248.1; -.
RT      HSSP: P17119; 3KAR.
DR      INTERPRO: IPR00253; -.
DR      INTERPRO: IPR001752; -.
DR      INTERPRO: IPR001753; -.
DR      PFAM: PF00225; kinesin.1.
DR      PFAM: PF00498; FHA.1.
DR      PRINTS: PR00380; KINESINHEAVY.
DR      PROSITE: PS00667; KINESIN_MOTOR_DOMAIN2.1.
SQ      SEQUENCE    1097 AA;  122333 MW;   8EF40B1C7579BA5B CRC64;

Query Match          59.2%; Score 1085.5; DB 11; Length 1097;
Best Local Similarity 60.5%; Pred. No. 1.6e-71;
Matches 216; Conservative 64; Mismatches 60; Indels 17; Gaps

QY      4  GGNIKVVVRVRRPNAREIDRGAKCIYRMEGNOTILTPPGAEEKARKSGKTIIDMGPKAF 63
       : ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      3  GASVKYAVRVRRPNARETSQDAKCVCVMGNTTISIINP-----KQSRMFL---KA-S 50

QY      64  FDRSYNSFDK-NAPNVARQEDLFQDLGVPLLDNAFKNYNCIFAYCGTSGCKSYSMGXG 122
       |||||
Db      51  FDYSYNSTSVSDPQFASQQOQVRYDGEEMLLHAFEGYNYCFAYGQTGACKSYTMWGRQ 110

QY      123  K--EHGVIPRIQCDFMRINELQKKNLTCTVEVSYLEIYNERVDLLNPSTKGNLKVRE 180
       : ::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      111  EPQGGGLVPOLCEDLFPSRVN-VNQSAQLSYSEVYSMEIICERVADLLNPKRSGLRVRE 169

QY      181  HPSTGPVFVEDLAKLVRSFOEINLMDECNKARTVAATNNNETSSRHAVFTTLTKQHW 240
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      170  HPILPGYPVDLSKLAVTSYADIADLMDCNKARTVAATNNNETSSRHAVFTIVTFQRSH 229
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF1B (FRAGMENT).
GN KIF1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Faure K., Gruber D., Bullinski J.C.;
RT "Identification of kinesin-like molecules in myogenic cells.";
RL Eur. J. Cell Biol. 0:0-0(1998).
DR EMBL; AF083331; AAC33292.1; -.
DR HSSP; P33176; 1BG2.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001356; -.
DR INTERPRO; IPR001752; -.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00027; HOMOBX_1; UNKNOWN_1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
FT NON_TER 689
SQ SEQUENCE 689 AA; 77352 MW; F64818C12F669319 CRC64;

Query Match 58.5%; Score 1073.5; DB 11; Length 689;
Best Local Similarity 60.2%; Pred. No. 6.1e-71;
Matches 216; Conservative 57; Mismatches 67; Indels 19; Gaps 7;
QY 4 GGNKKVVRVRFNAREIDRGAKIVMEGNQITLTTPPGAAEKARKSGKTMIDGPKAFA 63
DB 3 GASVAVVRVRFNAREIDRGAKIVMEGNQITLTTPPGAAEKARKSGKTMIDGPKAFA 63
QY 64 FDRSYWSP-DKNAPNYARQEDLFODLGVPLLDNAFKYNNCIFYAGTGSGKSYMMGYG 122
DB 52 FDSYWHTSPEDPCFASQSRVNDICKMLLHAFEGYNNCIFYAGTGSGKSYMMGYG 111
QY 123 KEH--GVIPRICDMFRINELQKDKNLTCTVEVSYLEIYNERVRLDNLNPS-TKGNLKVREHPSTG 180
DB 112 EESQAGIIPCCCELFKIND-NCNEDMSYSVS-SYMEIYCERVRDLNPKNKGNLKRVRE 169
QY 181 HPSYGPVDEDLAKLVRSFOEINLMDGKARTVAATNMNETSSRSHAVFTLTITLTKQW 240
DB 170 HPLGPIVEDLSKLVSYTDIADLADGAKARTVAATNMNETSSRSHAVFTIYVTKQK 229
QY 241 DEETKMDTEKVAKISLVDLGASERATSGATGARLKEGAEINRSLSLTGRVIAALADMS 300
DB 230 DPETNLSTEKVTKISLVDLGASERADSGATGARLKEGAEINRSLSLTGRVIAALAEVSK 289
QY 301 GKQKKNQVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETL--STLRYADSAK 357
DB 290 -KKKKTDFIPYRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADRAK 347

RESULT 11
ID 001349 PRELIMINARY; PRT; 1921 AA.
AC 001349;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE KINESIN-73.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97188425; PubMed=9037010;
RA Li H.P., Liu Z.M., Nirenberg M.;
RT "Kinesin-73 in the nervous system of Drosophila embryos.";

Proc. Natl. Acad. Sci. U.S.A. 94:1086-1091(1997).
DR EMBL; U81788; AAB50404.1; -.
DR HSSP; P33176; 1BG2.
DR FLYBASE; FBgn0019668; Khc-73.
DR INTERPRO; IPR000938; -.
DR INTERPRO; IPR001220; -.
DR INTERPRO; IPR001752; -.
DR PFAM; PF00225; kinesio; 1.
DR PFAM; PF01302; CAP_GLY; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00845; CAP_GLY; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1921 AA; 215047 MW; 4643F6F9783E99D0 CRC64;

Query Match 56.1%; Score 1028; DB 5; Length 1921;
Best Local Similarity 60.3%; Pred. No. 6.1e-67;
Matches 213; Conservative 45; Mismatches 85; Indels 10; Gaps 4;
QY 7 IKVVVRVRFNAREIDRGAKIVMEGNQITLTTPPGAAEKARKSGKTMIDGPKAFAFDR 66
DB 6 IKVAVRVRFNAREIDRGAKIVMEGNQITLTTPPGAAEKARKSGKTMIDGPKAFAFDR 58
QY 67 SYWDFDKNAPNYARQEDLFODLGVPLLDNAFKYNNCIFYAGTGSGKSYMMGYGKEHG 126
DB 59 CFYSLNPEDENFASQETVFCVGRGILDNAFQGYNACIFYAGTGSGKSYMMGYGKEHG 118
QY 127 VIPRICDMFRINELQKDKNLTCTVEVSYLEIYNERVRLDNLNPS-TKGNLKVREHPSTG 185
DB 119 IIPRICDQLFSAIAN-KSTPELTKYKVEVSYMEIYCERVRDLNPKNKGNLKRVREHVMG 177
QY 186 PYVEDLAKLVRSFOEINLMDGKARTVAATNMNETSSRSHAVFTLTITLTKQWDEETK 245
DB 178 PYVDGLSOLAVTSYQDIDNLMTGKNSRTVAATNMNAESSRSHAVSVLTILTDQATG 237
QY 246 MDEKVAKISLVDLGASERATSGATGARLKEGAEINRSLSLTGRVIAALADMSGKQK 305
DB 238 VSQKVSRLSLVDLAGSERAVTKGAVGRDLKEGSGNINSLTTLGLVSKLADQSGKSG 297
QY 306 N-OLVPRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADSAK 357
DB 298 NDKVPIRDSVLTWLLKDSLGGSMTAMIAAISPADINFEETLSTLRYADRAK 350

RESULT 12
ID 09V7C9 PRELIMINARY; PRT; 1921 AA.
AC 09V7C9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE KHC-73 PROTEIN.
GN KHC-73.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

```

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA De Padua B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang I., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-F., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003810; AAF58129.1; -.
DR HSP: P33176; 1962.
DR FLTBASE: FBgn0019968; Khc-73.
DR INTERPRO: IPR000938; -.
DR INTERPRO: IPR001220; -.
DR INTERPRO: IPR001752; -.
DR PFAM: PF00225; kinesin; 1.
DR PFAM: PF01302; CAP_GLY; 1.
DR PRINTS: PR00360; KINESINHEAVY.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS00845; CAP_GLY; 1.
DR PROSITE: PS00607; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE: PS00607; KINESIN_MOTOR_DOMAIN2; 1.
SQ SEQUENCE 1921 AA; 214997 MW; DEDCA75A63444C42 CRC64;

Query Match 56.1%; Score 1028; DB 5; Length 1921;
Best Local Similarity 60.3%; Pred. No. 61e-67;
Matches 213; Conservative 45; Mismatches 85; Indels 10; Gaps 4;

QY 7 IKVVVRPFPNAREIDRGAKCIVRMENGTITLTPPPGAEKARKSGKTMGPKAFADP 66
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6 IKVAVRPFPNRETEIDRCIVEMEKQOITLQNPPLPKIERKQ-----PKTEAFDH 58
::: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 67 SYNSFKNAPYARQEDFDGLVPLLDNAFKYNNCIFYAGTSGSKSYMMGCKEKG 126
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 59 CFYSLNPEDENFASQETVFCVCGILDNAGFYACIFYAGTSGSKSYMMGCKEKG 118
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 127 VIPRICODMFRINELQKRLCTVEVSYLEIYNERVRLDLPNS-TKGNLKVREHPSTG 185
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 119 IIPRLCDQLFSAIAN-KSTPELMKYVEVSYMEIYNEKVDHLLDPKPNQSLKVRHNVMG 177
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 186 PYVEDLAKLVRSFOETENLMDGNKARTVAATNNNETSSRSYHVFITLTQKWHDEETK 245
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 178 PYVDGLSGLVTSYQIDINLMTGNKSRVTAATNNNAESSRHVSVFVLTQITDQATG 237
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 246 MDEKVKAKLSLVDLGASERATSTGATGARKLKEGAEINRSLTSLGRVIAALDSSGKQK 305
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 238 VSGEKVRNSLVDLGASERAVKTCAGVGRLEKSGNINKSLTTLGLVLSKLADSNCKSG 297
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 306 N-QLVPRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSTLRYADSAK 357
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 298 NDKFVPRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSTLRYADRAK 350
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

```

```

RESULT 13
ID Q9NOT8 PRELIMINARY; PRT; 1826 AA.
AC Q9NOT8;
DT 01-OCT-2000 (TrEMBLrel. 15; Created)
DT 01-OCT-2000 (TrEMBLrel. 15; Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15; Last annotation update)
DE KINESIN-LIKE PROTEIN GAKIN.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Hanada T., Lin L., Tibaldi E.V., Reinherz E.L., Chishti A.H.;
RT "GAKIN: a novel kinesin-like protein associates with the human homolog
of the Drosophila discs large tumor suppressor in T lymphocytes.";
RL J. Biol. Chem. 0-0-0(2000).
DR EMBL: AF279865; AAF81263.1; -.
DR EMBL: AF279865; AAF81263.1; -.
SQ SEQUENCE 1826 AA; 202665 MW; C614E7F3A89E89ED CRC64;

Query Match 55.4%; Score 1016.5; DB 4; Length 1826;
Best Local Similarity 58.5%; Pred. No. 4e-66;
Matches 210; Conservative 49; Mismatches 87; Indels 13; Gaps 6;

QY 3 GGSNIKVVVRPFPNAREIDRGAKCIVRMENGTITLTPPPGAEKARKSGKTMGPKAF 62
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 2 GDSKVKVAVIRPWRNRETDLHTKCVVDVVDANKVILNPVNTLSKGDARGQ-----PKCF 56
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 63 AFDRSYWSFKNA-PNYARQEDFDGLVPLLDNAFKYNNCIFYAGTSGSKSYMMGY 121
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 57 AYDCHFWSDSESVKERYAGODIVFKLGENILQNAFDGYNACIFYAGTSGSKSYMMGT 116
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 122 GKEHVIPRICODMFRINELQKKN--LCTVEVSYLEIYNERVRLDLPNS-STGNLKV 178
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 117 ADQGLIPRLCSGLFER---TQKEENEQSFKEVSYMEIYNEKVRDLPKSGRQTLKV 173
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 179 RHPSTGYPVEDLAKLVRSFOETENLMDGNKARTVAATNNNETSSRSYHVFITLTQK 238
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 174 REHVLGYPYDGLSKLAATSYKDIESLSEGNKSRVTAATNNNEESSRHVITLTFT 233
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 239 WHDETQKMDTEKVKAKLSLVDLGASERATSTGATGARKLKEGAEINRSLTSLGRVIAALDM 298
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 234 LYDAKSGTSGCKVKLSLVDLGASERATKGAAGDLKSGNINESLTTLGLVISLADQ 293
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 299 SSGKOKNLQVPRDSVLTLLKDSLGNSMTAMIAAISPADINFEETLSTLRYADSAK 357
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 294 SAGK-NKNKFPYPRDSVLTLLKDSLGNSMTAMVATVSPADNVDYETLSTLRYADRAK 351
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 14
ID Q20888 PRELIMINARY; PRT; 1576 AA.
AC Q20888;
DT 01-NOV-1996 (TrEMBLrel. 01; Created)
DT 01-NOV-1996 (TrEMBLrel. 01; Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14; Last annotation update)
DE SIMILAR TO KINESIN-RELATED PROTEIN. NCBI GI: 1109842.
GN F56B3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:12:47 ; Search time 56.07 seconds
(without alignments)
186.588 million cell updates/sec

Title: US-09-235-416-1_COPY_602_784

Perfect score: 938

Sequence: 1 QEOSLRHSVTNSQLSPAP.....ELRQQQAQMEELAKTAKQEF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 350729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 350729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	938	100.0	784	20 Y06618	Thermomyces lanugi
2	102.5	10.9	436	21 G43544	Arabidopsis thalia
3	92.5	9.9	284	21 B43650	Human cancer assoc
4	91	9.7	592	21 G15484	Arabidopsis thalia
5	91	9.7	632	21 G46578	Arabidopsis thalia
6	91	9.7	644	21 G46577	Arabidopsis thalia
7	91	9.7	650	21 G15483	Arabidopsis thalia
8	91	9.7	702	21 G46576	Arabidopsis thalia
9	91	9.7	725	21 G15482	Arabidopsis thalia
10	90.5	9.6	326	21 G35664	Arabidopsis thalia
11	90.5	9.6	1816	21 Y95440	Caenorhabditis ele

12	89.5	9.5	590	20 Y41101	Rat gadd34 protein
13	89.5	9.5	1028	14 R41044	Plasmodium vivax D
14	89.5	9.5	1115	12 R13457	Duffy receptor, P
15	89.5	9.5	1115	18 W22478	Duffy antigen bind
16	89.5	9.5	1115	21 Y77899	P. vivax DABP bind
17	89.5	9.5	1245	16 R70106	TNF-R-PI. vivax Du
18	88	9.4	404	21 B19686	Rat NNG2 (RUKM1) P
19	88	9.4	427	21 B19687	Rat NNG2 (RUKM3) P
20	88	9.4	655	21 B19685	Rat NNG2 (RUK1) P
21	87.5	9.3	933	21 Y58435	Staphylococcus aur
22	87	9.3	476	22 B36590	Human FLEXH-12 pr
23	87	9.3	936	18 W89801	Staphylococcus aur
24	86.5	9.2	302	21 G35665	Arabidopsis thalia
25	86.5	9.2	340	19 W59552	Amino acid sequenc
26	86.5	9.2	623	21 G42595	Arabidopsis thalia
27	86.5	9.2	635	21 G42594	Arabidopsis thalia
28	86.5	9.2	693	21 G42593	Arabidopsis thalia
29	85.5	9.1	2100	20 W89579	Calcium permeable
30	84	9.0	653	21 Y94907	Human secreted pro
31	84	9.0	659	19 W37724	CD2 associated int
32	84	9.0	665	21 B19690	Human NNG2 (short
33	84	9.0	709	21 B19691	Human NNG2 (long f
34	83	8.8	1724	21 G31229	Arabidopsis thalia
35	83	8.8	1730	21 G31228	Arabidopsis thalia
36	83	8.8	1739	21 G50122	Arabidopsis thalia
37	83	8.8	1745	21 G50121	Arabidopsis thalia
38	83	8.8	1757	21 G31227	Arabidopsis thalia
39	83	8.8	1772	21 G50120	Arabidopsis thalia
40	82.5	8.8	257	21 G35780	Arabidopsis thalia
41	82.5	8.8	258	21 G35779	Arabidopsis thalia
42	82.5	8.8	1056	21 G31894	Arabidopsis thalia
43	82.5	8.8	1073	21 G31893	Arabidopsis thalia
44	82.5	8.8	1079	21 G48607	Arabidopsis thalia
45	82.5	8.8	1087	21 G31892	Arabidopsis thalia

ALIGNMENTS

RESULT	1
Y06618	
ID	Y06618 standard; Protein; 784 AA.
XX	
AC	Y06618;
XX	
DT	26-OCT-1999 (first entry)
XX	
DE	Thermomyces lanuginosus kinesin motor protein TL-gamma.
XX	
XX	TL-gamma; kinesin; motor protein; microtubule; unc-104; infection;
KW	neurodegenerative disease; Alzheimer's disease;
KW	Parkinson's disease; Huntington's disease;
KW	amyotrophic lateral sclerosis.
XX	
OS	Thermomyces lanuginosus.
XX	
PN	W09937659-A1.
XX	
PD	29-JUL-1999.
XX	
PF	22-JAN-1999; 99WO-US01355.
XX	
PR	23-JAN-1998; 98US-0072361.
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Goldstein LSB, Sakowicz R;
XX	
DR	WPI: 1999-493950/41.
DR	N-PSDB: X87656.
XX	
PT	New nucleic acid encoding microtubule motor protein, used for
	diagnosis of fungal infection and neurodegenerative disease

```
XX Claim 5; Page 70-71; 75pp; English.
PS
XX This sequence represents Thermomyces lanuginosus TL-gamma, a novel
CC ATP-dependent, plus end-directed microtubule motor protein that is
CC a member of the unc-104 family and kinesin superfamily. The
CC invention provides TL-gamma nucleic acids (see X87656), proteins
CC and antibodies, and methods of screening for TL-gamma modulators
CC potentially useful for treating hyphal fungal infections and
CC diseases caused by mutated TL-gamma, e.g. neurodegeneration
CC involving anterograde axonal transport, such as Alzheimer's,
CC Parkinson's or Huntington's diseases or amyotrophic lateral
CC sclerosis. Detection of TL-gamma allows differentiation between
CC hyphal and non-hyphal fungal infections.
XX
SQ Sequence 784 AA;

Query Match 100.0%; Score 938; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.6e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEQSLRHVNTNSQLGSPAGRHRTLSKAGSDAGDSRSDSPLPHFRGKDSDFYARRE 60
Db |||
QY 602 qeqslrhsvtnsqgspagrhrtlskagsdagsdsplphfrgkdsdfyarre 661
Db |||
QY 61 AASAILGLDOKISHLTDDELDAFDVQKARAVRRLGLEDNEDSDSSPPVRDKYMSNG 120
Db |||
QY 662 aasailglqkshltddeldalfdvqkaravrglvedneddsqssfpvrckymnsng 721
Db |||
QY 121 TIDNFSLDTAITMPTGPRSDDDGDALEFFGDKKSKQDASNVDDLELROOQAOQMEALKTAK 180
Db |||
QY 722 tidnfsldtaitmptgprsdddgdalffgdkkksqdasnvddlelrrqqqameealktak 781
QY 181 QEF 183
Db |||
QY 782 qef 784

RESULT 2
G43544
ID G43544 standard; Protein; 436 AA.
XX
AC G43544;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54439.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123348.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139452.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
```

PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 10.9%; Score 102.5; DB 21; Length 436;		
Best Local Similarity 24.68; Pred No. 0.035;		
Matches 28; Conservative 26; Mismatches 53; Indels 7; Gaps		
QY	69	DKTSHTTDELDALFDVQKARAVRGLVEDNDSQSFVPRDKYMSNGTIDNFSLD 128
Db	35	dqlpgldicdevdsmedgeseeeddegdtteeddegdsceddegenkedgedesedfe-- 92
		: :: : : :: : : : : : : : : : : : : : : :
QY	129	TAITWPTGTPRSSDGDALFFGDKSKODASNYDVELRQQAOMEALKTAQOE 182
Db	93	-----dgnk'esegsdgndnkdqameelekev'elrsqeqdilknrkdkge 141
		: :
RESULT	3	
ID	B43650	
AC	XX	B43650 standard; Protein; 264 AA.
AC	XX	B43650;
DT	08-FEB-2001	(first entry)
DE	XX	Human cancer associated protein sequence SEQ ID NO:1095.
KW	XX	Human; cancer associated gene; cancer antigen; detection; cancer;
KW	XX	diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
KW	XX	antidiabetic; antilasthmatic; antirheumatic; antiallergic; antiviral;
KW	XX	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW	XX	dermatological; neuroprotective; thrombolytic; coagulant; neurologic;
KW	XX	vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW	XX	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	XX	allergic reaction; graft versus host disease; organ rejection;
KW	XX	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	XX	neurological disease; drug screening.
OS	XX	Homo sapiens.
PN	XX	WO200055350-A1.
PD	XX	21-SEP-2000.
PF	XX	08-MAR-2000; 2000WO-US05882.
PR	XX	12-MAR-1999; 99US-0124270.
PA	XX	(HUMA-) HUMAN GENOME SCI INC.
PI	XX	Rosen CA, Ruben SM;
DR	XX	WPI; 2000-587533/55.
DR	XX	N-PSDB; C77859.

PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 03-AUG-1999; 99US-0145389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0153363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0155658.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.7%; Score 91; DB 21; Length 592;
Best Local Similarity 23.9%; Pred. No. 0.77; Mismatches 30; Indels 38; Gaps 9;
Matches 45; Conservative 30; Mismatches 30; Indels 38; Gaps 9;
QY 21 GRHDTLSKAGSDADGSRSDSLPHFRGKSDW-----FYARREANSAIIGLDOKISHL 75
Db 259 yghgssksatkpspaprpsppppssgesdsggplgafwatqhaktavssednk---- 314
QY 76 TDDELALPDYQKARVRRGLVEDNEDSDSQSPFVRDKYMSNGTIDNFSLDTAITMPG 135
Db 315 mmpkfdepsntskservr---vdshq---pkpsspvrye--arglqrnkdtettisqkn 366
QY 136 TP-----RSDD-----DGDALFF--GDKKSKQDASNDVVEELROO--QAQME 174
Db 367 ttpaaanmtryskddafnsfvadfttkfdngnkpgeealeaeiqrkdelkqtksk 426
QY 175 ALKTAQOE 182
Db 427 aeitakfe 434
RESULT 5
G46578
ID G46578 standard; Protein; 632 AA.
XX
AC G46578;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58613.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 05-MAY-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 9.7%; Score 91; DB 21; Length 632;		
Best Local Similarity 23.9%; Pred No. 0.84; Mismatches 30; Conservative 30; Indels 38; Gaps 9;		
Oy	21	GRHRTLSKAGSDGSDSPHFRGKDSW-----FYARRAASAILGLDQKISHL 75
Db	247	gvhdgssksatkspapr:spppppsgesdgpglgaftatqhtaktsvvsednk---- 302
Oy	76	TDELDALFDVQKARAVRGLVEDNEDSDSQSFVPRDKYMSNGTIDNFSLDTAITMPG 135
Db	303	nmpkdepnstkservr---vdshq---pkp:spvrge--arg:qrnkldettisqkn 354
Oy	136	TP-----RSD-----DGDALFF--GDKXSKODASNVDELRRQ--QAQME 174
Db	355	tptaaanmtrvskddafnsfvadfttkfdngnkp:gealeaigrlkdeliktksk 414
Oy	175	ALKTARQE 182
Db	415	aeltakfe 422
RESULT 6		
ID	G46577	
XX	AC	G46577;
XX	DT	18-OCT-2000 (first entry)
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58612.
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;

KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence..
XX	Arabidopsis thaliana.
XX	EP1033405-A2.
XX	06-SEP-2000.
PF	25-FEB-2000; 2000EP-0301439.
XX	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.
PR	21-JUN-1999; 99US-0139817.
PR	22-JUN-1999; 99US-0139899.
PR	23-JUN-1999; 99US-0140353.
PR	23-JUN-1999; 99US-0140354.

DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 15752.	
DE	Protein identification; signal transduction pathway; metabolic pathway;	
XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
KW	Arabidopsis thaliana.	
XX	EPI033405-A2.	
XX	06-SEP-2000.	
PD		
XX	25-FEB-2000; 2000EP-0301439.	
PF		
XX	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	28-APR-1999; 99US-0130891.	
PR	30-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	
PR	04-MAY-1999; 99US-0132407.	
PR	04-MAY-1999; 99US-0132484.	
PR	05-MAY-1999; 99US-0132485.	
PR	06-MAY-1999; 99US-0132486.	
PR	07-MAY-1999; 99US-0132487.	
PR	11-MAY-1999; 99US-0132863.	
PR	14-MAY-1999; 99US-0134256.	
PR	14-MAY-1999; 99US-0134218.	
PR	14-MAY-1999; 99US-0134219.	
PR	14-MAY-1999; 99US-0134221.	
PR	14-MAY-1999; 99US-0134370.	
PR	18-MAY-1999; 99US-0134768.	
PR	19-MAY-1999; 99US-0134941.	
PR	20-MAY-1999; 99US-0135124.	
PR	21-MAY-1999; 99US-0135353.	
PR	24-MAY-1999; 99US-0135629.	
PR	25-MAY-1999; 99US-0136021.	
PR	27-MAY-1999; 99US-0136392.	
PR	28-MAY-1999; 99US-0136782.	
PR	01-JUN-1999; 99US-0137222.	
PR	03-JUN-1999; 99US-0137528.	
PR	04-JUN-1999; 99US-0137502.	
PR	07-JUN-1999; 99US-0137724.	
PR	08-JUN-1999; 99US-0138094.	
PR	10-JUN-1999; 99US-0138540.	
PR	10-JUN-1999; 99US-0138847.	
PR	14-JUN-1999; 99US-0139119.	
PR	16-JUN-1999; 99US-0139452.	
PR	16-JUN-1999; 99US-0139453.	
PR	17-JUN-1999; 99US-0139492.	
PR	18-JUN-1999; 99US-0139454.	
PR	18-JUN-1999; 99US-0139455.	
PR	18-JUN-1999; 99US-0139456.	
PR	18-JUN-1999; 99US-0139457.	
PR	18-JUN-1999; 99US-0139458.	
PR	18-JUN-1999; 99US-0139459.	
PR	18-JUN-1999; 99US-0139460.	
PR	18-JUN-1999; 99US-0139461.	
PR	18-JUN-1999; 99US-0139462.	
PR	18-JUN-1999; 99US-0139463.	
PR	18-JUN-1999; 99US-0139750.	

PR	18-JUN-1999; 99US-0139763.	
PR	21-JUN-1999; 99US-0139817.	
PR	22-JUN-1999; 99US-0139899.	
PR	23-JUN-1999; 99US-0140353.	
PR	23-JUN-1999; 99US-0140354.	
PR	24-JUN-1999; 99US-0140695.	
PR	28-JUN-1999; 99US-0140823.	
PR	29-JUN-1999; 99US-0140991.	
PR	30-JUN-1999; 99US-0141287.	
PR	01-JUL-1999; 99US-0141842.	
PR	01-JUL-1999; 99US-0142154.	
PR	02-JUL-1999; 99US-0142055.	
PR	06-JUL-1999; 99US-0142380.	
PR	08-JUL-1999; 99US-0142803.	
PR	09-JUL-1999; 99US-0142920.	
PR	12-JUL-1999; 99US-0142977.	
PR	13-JUL-1999; 99US-0143542.	
PR	14-JUL-1999; 99US-0143624.	
PR	15-JUL-1999; 99US-0144005.	
PR	16-JUL-1999; 99US-0144085.	
PR	16-JUL-1999; 99US-0144086.	
PR	19-JUL-1999; 99US-0144325.	
PR	19-JUL-1999; 99US-0144331.	
PR	19-JUL-1999; 99US-0144332.	
PR	19-JUL-1999; 99US-0144333.	
PR	19-JUL-1999; 99US-0144334.	
PR	19-JUL-1999; 99US-0144335.	
PR	20-JUL-1999; 99US-0144352.	
PR	20-JUL-1999; 99US-0144632.	
PR	20-JUL-1999; 99US-0144884.	
PR	21-JUL-1999; 99US-0144884.	
PR	21-JUL-1999; 99US-0145086.	
PR	21-JUL-1999; 99US-0145088.	
PR	22-JUL-1999; 99US-0145085.	
PR	22-JUL-1999; 99US-0145087.	
PR	22-JUL-1999; 99US-0145089.	
PR	22-JUL-1999; 99US-0145192.	
PR	23-JUL-1999; 99US-0145145.	
PR	23-JUL-1999; 99US-0145218.	
PR	23-JUL-1999; 99US-0145224.	
PR	26-JUL-1999; 99US-0145276.	
PR	27-JUL-1999; 99US-0145913.	
PR	27-JUL-1999; 99US-0145918.	
PR	27-JUL-1999; 99US-0145919.	
PR	28-JUL-1999; 99US-0145951.	
PR	02-AUG-1999; 99US-0146386.	
PR	02-AUG-1999; 99US-0146388.	
PR	02-AUG-1999; 99US-0146389.	
PR	03-AUG-1999; 99US-0147038.	
PR	04-AUG-1999; 99US-0147204.	
PR	04-AUG-1999; 99US-0147302.	
PR	05-AUG-1999; 99US-0147192.	
PR	05-AUG-1999; 99US-0147260.	
PR	06-AUG-1999; 99US-0147303.	
PR	06-AUG-1999; 99US-0147416.	
PR	09-AUG-1999; 99US-0147493.	
PR	09-AUG-1999; 99US-0147935.	
PR	10-AUG-1999; 99US-0148171.	
PR	11-AUG-1999; 99US-0148319.	
PR	12-AUG-1999; 99US-0148341.	
PR	13-AUG-1999; 99US-0148565.	
PR	13-AUG-1999; 99US-0148684.	
PR	16-AUG-1999; 99US-0149368.	
PR	17-AUG-1999; 99US-0149175.	
PR	18-AUG-1999; 99US-0149426.	
PR	20-AUG-1999; 99US-0149722.	
PR	20-AUG-1999; 99US-0149723.	
PR	20-AUG-1999; 99US-0149929.	
PR	23-AUG-1999; 99US-0149902.	
PR	23-AUG-1999; 99US-0149930.	
PR	25-AUG-1999; 99US-0150566.	
PR	26-AUG-1999; 99US-0150884.	
PR	27-AUG-1999; 99US-0151065.	

```
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 23.9%; Pred. No. 0.87;
Matches 45; Conservative 30; Mismatches 75; Indels 38; Gaps 9;

QY 21 GRDRTLKAGSDADGDSRSDSLPHERKXDSW-----PYAREASAILGLDOKISHL 75
Db 317 gvhdgssksatkpspprrppppssgsdsggplgarfwatghaktvsvednk---- 372
QY 76 TDDELDAFDVQKARAVRGLVEDNEDSDSQSFVPVRDKYMSNGTIDNFSLDTAITMPG 135
Db 373 nmpkfdepnatskservr---vdshq---pkpkspvurge--arglqrnkdllettisqkn 424
QY 136 TP-----RSD-----DGDALFF--GDKKSKQDASNVVDVEELRQ--QAQME 174
Db 425 ttpaaannmrsvskddafnsfvadfttkfdngnkpqkkealeaeigrkldelkqtkek 484
QY 175 ALKTAKOE 182
Db 485 aeatkre 492

RESULT 8
```

```
G46576
ID G46576 standard; Protein; 702 AA.
XX
AC G46576;
DT
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58611.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
```



```
Db      485 aeltakfe 492
RESULT  9
GL15482
ID      G15482 standard; Protein; 725 AA.
XX
XX
AC      G15482;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 15751.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
XX      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD
PF      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      08-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      23-APR-1999; 99US-0130891.
PR      28-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132484.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139753.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
```


OY 58 RREASAILGLD--QKISHLTDELDALPDVQKARAVRGLVEDNEDSDQSFPVRDK 115
Db 74 -----sgmlvqdpckeph---ddvdeifnqarlgavegpl-----ehpssrsftgtgr 121
OY 116 YMSGTIDNFSLDTAITMP-----GTPRSDDDGDALFFGDKKSKODA 157
Db 122 llage-----avptalqgpepvihniifwsngftvddgplrlkldpenasfldsirksec 176
OY 158 SNVDVEEL-----ROOQAQMEALKTAQ 181
Db 177 pk-elepvdkrapvhnlmrrdckcekekksfsfq 211
RESULT 11
Y95440
ID Y95440 standard; Protein; 1816 AA.
AC Y95440;
XX
DT 10-OCT-2000 (first entry)
XX
DE Caenorhabditis elegans polypeptide at the c05c12.3 locus.
XX
KW SOC/CRAC; calcium channel; store operated channel;
KW calcium release activated channel; therapy; diagnosis;
KW lymphocyte proliferative disorder.
XX
OS Caenorhabditis elegans.
XX
PN W0200040614-A2.
XX
PD 13-JUL-2000.
XX
PF 20-DEC-1999; 99WO-US29996.
XX
PR 30-DEC-1998; 98US-0114220.
PR 29-JAN-1999; 99US-0120018.
PR 22-JUN-1999; 99US-0140415.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX Scharenberg AM;
XX WPI; 2000-465957/40.
XX
PT New SOC/CRAC calcium channel polynucleotides and polypeptides used to
PT diagnose and treat proliferative disorders associated with the channel,
PT and to screen for novel modulators of the channel -
XX
PS Example; Page 72-76; 108pp; English.
XX
CC The present sequence is that of a Caenorhabditis elegans
CC polypeptide at the c05c12.3 locus. The polypeptide was identified
CC in a database search for putative calcium channel proteins. The
CC C05C12.3 protein was notable cause its central pore region had
CC some similarity to, but was clearly distinct from, members of the
CC Trp family of calcium channels. The polypeptide was used in
CC BLAST screening to isolate 2 other C. elegans homologues (see
CC Y95441-42) and mouse melastatin-1 (see Y95438). These were used
CC to screen EST databases for lymphocyte homologues. Human clones
CC (see A49923-24) encoding members (see Y95435-37) of a new family of
CC SOC (store operated channel) or CRAC (calcium release activated
CC channel) calcium channel polypeptides were identified.
XX
SQ Sequence 1816 AA;
Query Match 9.6%; Score 90.5; DB 21; Length 1816;
Best Local Similarity 21.9%; Pred. No. 3.9;
Matches 48; Conservative 30; Mismatches 62; Indels 79; Gaps 11;
OY 12 NSQLGSPAPGRHNR-TLSKAGSDAGDSRS-----DSPLPFRGKSDWF 55

Db 1610 nkklgkflsghkyrgsfsgsgsdggsgssdnslkpnnsvpmitvdgppp-igsrrtsagq 1668
OY 56 YARREASAILGLDQOKISHLTDELDALFDVQKARAVRGLVEDNEDSDQS----- 108
Db 1669 ylkrdslqa-----kkki---tenrrsal--eqkipsiqfmlmedgdaaesateevsi 1719
OY 109 SPV-----RDK-----YMSG-----TIDN 124
Db 1720 sipvpqmrvrqvtesksdlseddllitredaptsiniprrrhalystiadaletedd 1779
OY 125 FSLDTATMPTGTPRSDDDGDALFFGDKKSKQDASVNDVB 163
Db 1780 fyadspvpmpmtpvqpadgs--figendsryqrddsyd 1816
RESULT 12
Y41101
ID Y41101 standard; Protein; 590 AA.
AC Y41101;
XX
DT 17-JAN-2000 (first entry)
XX
DE Rat gadd34 protein sequence.
XX
KW Progression-elevated gene-3; PEG-3 gene; regulatory region; cell death;
KW cell growth; cancer; thymidine kinase; gancyclovir; acyclovir; rat;
KW gadd34 protein.
XX
OS Rattus sp.
XX
PN W09949898-A1.
XX
PD 07-OCT-1999.
XX
PF 31-MAR-1999; 99WO-US07199.
XX
PR 31-MAR-1998; 98US-0052753.
XX
XX (UYCO) UNIV COLUMBIA NEW YORK.
XX Fisher PB;
XX WPI; 1999-591184/50.
XX
PT Novel vectors useful for studying the progression of cancer -
PS Disclosure; Fig 3A-B; 251pp; English.
XX
CC The invention relates to an inducible progression-elevated gene-3 (PEG-3
CC gene) regulatory region functionally linked to a gene encoding a product
CC that causes or may be induced to cause the death or inhibition of cancer
CC cell growth. A vector of the invention which contains a gene encoding
CC thymidine kinase or a product which causes the cell to express a
CC specific antibody can be administered along with gancyclovir or acyclovir,
CC or an antibody or fragment to the antigen, respectively, to treat cancer
CC in a subject. The PEG-3 gene is useful for generating new cloning and
CC expression vectors, transfected cells, and for developing methods for
CC cultured growth of such cells. The PEG-3 polynucleotide is also useful as
CC a source of primers and probes to study the progression of cancer, and to
CC detect the presence of the gene. The present sequence represents the
CC amino acid sequence of the rat gadd34 protein.
XX
SQ Sequence 590 AA;
Query Match 9.5%; Score 89.5; DB 20; Length 590;
Best Local Similarity 25.7%; Pred. No. 1.1;
Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;
OY 15 LGSFAPGRHD---RTLSKAGSDAGDSRSPLPFRGKSDWTFYARREASAILGLDOK 71

Db 309 lssptspehdfkawyrgedteddd-----dsdwgsaeeg-----ka 348
 QY 72 ISHLTDDELDFDQKARAVRGLVEDNEDSQSFFPVRDKYMSNGTIDNFS----- 126
 Db 349 lssptspehdf-----kawyrpg--edited-dqdsdwgsaek---dglafatphts 397
 QY 127 --LDTAITMPTPRSDGDDGALFFGDKKSKQDASNVVDELQQQAQ 171
 Db 398 afiktwccpgedtedddcevvv-----pedseaadpdkspsheaq 438

RESULT 13
 R41044
 ID R41044 standard; protein; 1028 AA.
 AC R41044;
 XX
 DT 22-MAR-1994 (first entry)
 XX Plasmodium vivax Duffy receptor residues 23-1051.
 DE Merozoite; malaria; HIV; env; CD4; truncated Duffy receptor;
 KW human immunodeficiency virus; envelope glycoprotein; hybrid protein;
 KW red blood cell; erythrocyte; AIDS; cross-linking agent;
 KW joining region.
 XX Plasmodium vivax.
 OS Plasmodium vivax.

Key Location/Qualifiers
 Modified-site 1028
 /note= "opt. terminates with a Cys residue for disulphide bond formation with an added Cys at the C-terminal end of a CD4 fragment; alternatively, Cys residues are incorporated near the C-terminal end of P.vivax Duffy 23-1051 and of CD4 1-371 to allow disulphide cross-linking or the two parts of the fusion protein could be connected via any linking molecule or agent"

WO9318160-A.
 XX
 XX 16-SEP-1993.
 XX
 XX 10-MAR-1993; 93WO-GB00505.
 XX
 XX 11-MAR-1992; 92GB-0005276.
 PR 08-JUL-1992; 92GB-0014481.
 PR 24-JUL-1992; 92GB-0015829.
 PR 16-SEP-1992; 92GB-0019562.
 PR 03-MAR-1993; 93GB-0004311.
 XX
 PA (PREN/) PRENDERGAST K F.
 XX
 XX Prendergast KF;
 PI
 XX WPI; 1993-303474/39.
 DR
 XX
 XX Anti-viral fusion peptide(s) - comprise viral-binding component
 PT and malaria merozoite red cell binding component, for treating
 PT e.g. HIV, and hepatitis
 PT
 XX
 PS Claim 11; Page 49-51; 69pp; English.
 XX
 CC A hybrid protein in which the P.vivax Duffy receptor is joined to
 CC the CD4 receptor molecule, both molecules being truncated at their
 CC transmembrane domains, is specifically claimed. The fusing or
 CC joining of the two segments takes place by joining amino acid
 CC segments at unspecified sites by disulphide bonds or by cross-linking
 CC agents of any type. The fusion protein can bind free HIV in the blood
 CC to red blood cells and consequently reduce viral titre, prevent
 CC transmission of the virus and improve safety of blood transfusions.
 CC The specification also includes the sequence of amino acids 1-371

CC of CD4 (see e.g. R41041).
 XX
 SQ Sequence 1028 AA;
 Query Match 9.5%; Score 89.5; DB 14; Length 1028;
 Best Local Similarity 24.0%; Pred. No. 2.3;
 Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;
 QY 6 LRHSVNSQLGSPAPGRHRTLSKAGSDADGDSPLPHFRGKSDWPFYARREASAI 65
 Db 694 lrhskdnssdsgpa-----esmanpdsnskgetg----kgqndmakatkdsnss 740
 QY 66 LGLDQKISHLTDDELDFDQKARAVRGLVEDNEDSQSFFPVRDKYMSN--GTI 122
 Db 741 dg-----tssatgdttdav-----dreinkgvped-----rdktvsgkdggge 778
 QY 123 DNFSLDTAITMPTPRSDGDDGALFFGDKKSKQDASNVVDEE---LRQOQAQMEALKTA 179
 Db 779 dmsankdaatvvgedrirensag-----gstndrskndtekgastpdksgsdals 832
 QY 180 QKE 182
 Db 833 kte 835

RESULT 14
 R13457
 ID R13457 standard; Protein; 1115 AA.
 XX
 AC R13457;
 DT 07-NOV-1991 (first entry)
 XX Duffy receptor.
 DE Malaria; parasite; vaccine.
 KW Plasmodium vivax Salvador I strain.
 XX
 FH Key Location/Qualifiers
 FT Region 1..1050
 FT /label= exon 1
 FT Region 1051..1077
 FT /label= exon 2
 FT Region 1078..1101
 FT /label= exon 3
 FT Region 1102..1115
 FT /label= exon 4
 FT Peptide 1..22
 FT /label= signal sequence
 XX
 XX US7554837-A.
 PN
 XX
 XX 23-JUL-1991.
 PD
 XX
 XX 20-JUL-1990; 90US-0554567.
 PF
 XX
 XX 20-JUL-1990; 90US-0554837.
 PR
 XX
 XX (USSH) NAT INST OF HEALTH.
 PA
 XX
 PI Miller L, Adams J, Kaslow D, Xiangdong F;
 XX
 XX WPI; 1991-260184/35.
 DR
 XX N-PSDB; Q13317.
 XX Plasmodium Duffy receptor proteins - used in vaccines against
 PT malaria and for producing antibodies which prevent malaria
 PT infection.
 PT
 XX Disclosure; Fig 12; 63pp; English.
 PS
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:13:24 ; Search time 31.19 Seconds
(without alignments)
112.715 Million cell updates/sec

Title: US-09-235-416-1_COPY_602_784

Perfect score: 938

Sequence: 1 OEOALLRHVTSNLSGSPAP.....ELRQQAQMEELAKTAKQEF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	89.5	9.5	590	3	US-08-893-852A-4	Sequence 4, Appl
2	89.5	9.5	590	4	US-08-821-818-2	Sequence 2, Appl
3	89.5	9.5	1115	2	US-08-568-459A-2	Sequence 2, Appl
4	89.5	9.5	1115	2	US-08-487-826B-2	Sequence 2, Appl
5	89.5	9.5	1115	6	5198347-6	Patent No. 5198347
6	87.5	9.3	933	3	US-08-293-728-2	Sequence 2, Appl
7	87.5	9.3	933	4	US-09-421-868-2	Sequence 2, Appl
8	85.5	9.1	1820	3	US-07-988-289B-8	Sequence 8, Appl
9	85.5	9.1	2100	2	US-08-808-793-23	Sequence 23, Appl
10	85.5	9.1	2100	3	US-08-772-512A-19	Sequence 19, Appl
11	80	8.5	1087	1	US-08-264-002-5	Sequence 5, Appl
12	78	8.3	3052	2	US-08-557-122A-26	Sequence 26, Appl
13	76.5	8.2	200	2	US-08-531-525-33	Sequence 33, Appl
14	76.5	8.2	200	2	US-08-718-270A-33	Sequence 33, Appl
15	76.5	8.2	1503	4	US-08-976-255-14	Sequence 14, Appl
16	75.5	8.0	290	2	US-08-903-801-1	Sequence 1, Appl
17	75.5	8.0	780	1	US-08-485-621-2	Sequence 2, Appl
18	75.5	8.0	780	2	US-08-973-831-2	Sequence 2, Appl
19	75.5	8.0	780	3	PCT-US96-09530A-2	Sequence 2, Appl
20	75	8.0	454	3	US-08-348-518C-4	Sequence 4, Appl
21	75	8.0	454	3	US-08-476-509B-4	Sequence 4, Appl
22	74	7.9	2105	2	US-08-808-793-3	Sequence 3, Appl
23	74	7.9	2105	3	US-08-772-512A-3	Sequence 3, Appl
24	73.5	7.8	204	2	US-08-531-525-32	Sequence 32, Appl
25	73.5	7.8	204	2	US-08-718-270A-32	Sequence 32, Appl
26	73.5	7.8	889	5	PCT-US93-11725-2	Sequence 2, Appl
27	73.5	7.8	4550	2	US-08-804-227C-8	Sequence 8, Appl

28	73.5	7.8	4550	2	US-08-804-198-2	Sequence 2, Appl
29	72.5	7.7	326	2	US-08-757-653-172	Sequence 172, App
30	72.5	7.7	326	2	US-08-823-516-75	Sequence 75, Appl
31	72.5	7.7	326	2	US-08-823-516-135	Sequence 135, App
32	72.5	7.7	326	3	US-08-759-038-111	Sequence 111, App
33	72.5	7.7	326	3	US-08-758-314-111	Sequence 111, App
34	72.5	7.7	380	2	US-09-026-587-4	Sequence 4, Appl
35	72.5	7.7	380	2	US-09-227-420-4	Sequence 4, Appl
36	72	7.7	764	1	US-08-375-300-4	Sequence 4, Appl
37	72	7.7	764	3	US-09-177-431-4	Sequence 4, Appl
38	72	7.7	764	5	PCT-US95-16930-4	Sequence 4, Appl
39	72	7.7	1089	1	US-08-375-300-2	Sequence 2, Appl
40	72	7.7	1089	3	US-09-177-431-2	Sequence 2, Appl
41	72	7.7	1089	5	PCT-US95-16930-2	Sequence 2, Appl
42	72	7.7	1349	4	US-08-938-291A-6	Sequence 6, Appl
43	71.5	7.6	249	1	US-08-466-603-2	Sequence 2, Appl
44	71.5	7.6	249	1	US-08-314-503A-2	Sequence 2, Appl
45	71.5	7.6	249	1	US-08-468-066-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-893-852A-4
; Sequence 4, Application US/08893852A
; Patent No. 6080558
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW HUMAN GROWTH REGULATOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,852A
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0341 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 452490
; US-08-893-852A-4

Query Match 9.5%; Score 89.5; DB 3; Length 590;
Best Local Similarity 25.7%; Pred. No. 0.15;

Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;

QY 15 LGSPAPGRHD---RTLKAGSDADGDSRSDSPLPFRGKDSDFYARREASAILGLDQK 71
Db 309 LSSPTSPEHDFLKAWYRPGEDTDD-----DSDWGSAAEEG-----KA 348
QY 72 ISHLTDELDAFDVQKARAVRGLVEDNEDSDSQSFFVRDKYMSNGTIDNFS----- 126
Db 349 LSSPTSPEHDFL-----KAWYRPG---EDTED-DQSDWGSAAEK---DGLAQTFATPHTS 397
QY 127 --LDTATMPGTPTSRDDGDALFFGDKKSKQDASNVVDVEELRQOQAA 171
Db 398 AFLKTWVCCPGEDTDDDCVWV-----PEDSEADPDKSPSHEAQ 438

RESULT 2

US-08-821-818-2
; Sequence 2, Application US/08821818
; Patent No. 6146877
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; TITLE OF INVENTION: IDENTIFICATION OF THE PROGRESSION
; TITLE OF INVENTION: ELEVATED GENE-3 AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,818
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Chan, Albert Wai-Kit
; REGISTRATION NUMBER: 36,479
; REFERENCE/DOCKET NUMBER: 51523
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-821-818-2

Query Match 9.5%; Score 89.5; DB 4; Length 590;
Best Local Similarity 25.7%; Pred. No. 0.15;
Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;
QY 15 LGSPAPGRHD---RTLKAGSDADGDSRSDSPLPFRGKDSDFYARREASAILGLDQK 71
Db 309 LSSPTSPEHDFLKAWYRPGEDTDD-----DSDWGSAAEEG-----KA 348
QY 72 ISHLTDELDAFDVQKARAVRGLVEDNEDSDSQSFFVRDKYMSNGTIDNFS----- 126
Db 349 LSSPTSPEHDFL-----KAWYRPG---EDTED-DQSDWGSAAEK---DGLAQTFATPHTS 397
QY 127 --LDTATMPGTPTSRDDGDALFFGDKKSKQDASNVVDVEELRQOQAA 171
Db 398 AFLKTWVCCPGEDTDDDCVWV-----PEDSEADPDKSPSHEAQ 438

RESULT 3

US-08-568-459A-2
; Sequence 2, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1115 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
US-08-568-459A-2

Query Match 9.5%; Score 89.5; DB 2; Length 1115;
Best Local Similarity 24.0%; Pred. No. 0.39;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;
QY 6 LRHSVTNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPFRGKDSDFYARREASAI 65
Db 716 LRHSKDNSSDSDGPA-----ESMANPDSNSKGTG---KGQNDMAKATKDSNSS 762
QY 66 LGLDQKTSHTLTDDELDAFDVQKARAVRGLVEDNEDSDSQSFFVRDKYMSN---GTI 122
Db 763 DG-----TSSATGDTTDAV-----DREINKGVPE-----RDKTVGSKDGGG 800
QY 123 DNFSLDTAITMPGTPTSRDDGDALFFGDKKSKQDASNVVDVEE---LRQOQAQMEALKTA 179
Db 801 DNSANKDAATVVGEDRIENSAG-----GSTNDRSKNDTEKNGCASTPDSKQSDATALS 854
QY 180 QKE 182
Db 855 KTE 857
RESULT 4


```

: Sequence 2, Application US/09421868-2
: Patent No. 6177084
: GENERAL INFORMATION:
: APPLICANT: Foster, Timothy J.
: APPLICANT: McDevitt, Damien L.
: TITLE OF INVENTION: The S. aureus
: FILE REFERENCE: 05344.105011
: CURRENT APPLICATION NUMBER: US/09/
: CURRENT FILING DATE: 1999-10-19
: PRIOR APPLICATION NUMBER: 08/293,
: PRIOR FILING DATE: 1994-08-22
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatencIn ver. 2.0
: SEQ ID NO 2
: LENGTH: 933
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
US-09-421-868-2

```

			9.3%;	Score 87.5;	DB 4;	Length 933;
			Best Local Similarity 25.6%;	Pred. No. 0.49;		
			Matches 42;	Conservative 18;	Mismatches 61;	Indels 43; Gaps 5;
QY	9 SVTNSQLGSPAPGRHRTLSKAGSDADGDSRDSPLPHFRCKDSDWFFARREASAILGL 68	: : :	:	:	:	:
Dd	795 SDSDDSDSDSDSESD---SDESDDSDSDSDSDSDSDSDSD-----SASDSDSGS 845	:	:	:	:	:
QY	69 DQKTSHLTDDDLDALEFDVQKARVRRLGVEDNEDDSQSFSF----PYRDKYMSNGTTIDN 124	:	:	:	:	:
Dd	846 DSDSSDSDSESDS-----NDSSESGSNNVVPPNPSPKMGTNASKNK 887	:	:	:	:	:
QY	125 FSLDTATMCPGTTPSRDDD-----GDALFFDGKKSKQD 156	:	:	:	:	:
Dd	888 EAKDSKEPLPTGTGEDANTSLLIWLGIASIGSLLFRRRKKNKD 931	:	:	:	:	:

```

RESULT      8
US-07-998-2898-B
; Sequence 8, Application US/07998289B
; Patent No. 6027876
; GENERAL INFORMATION:
; APPLICANT: Black, Bruce C
; APPLICANT: Taylor, Martin
; APPLICANT: Heckel, David G
; TITLE OF INVENTION: Method for Monitoring
; TITLE OF INVENTION: Resistance
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Ve
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,289B
; FILING DATE: 30-DEC-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/0A939
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 8:

```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-998-289B-8

```

Query Match 9.1%; Score 85.5; DB 3; Length 1820;
Best Local Similarity 21.9%; Pred. NO.2.1;
Matches 44; Conservative 36; Mismatches 70; Indels 51; Gaps 9;

Qy	4	SLLRHVTNSQLSGSPAGRHDRITLKSAGSDADGDSRSUSPLPHFRGKDSDFYARRAAS	63
Db	1016	ALLSNFGSSLSAPT-----ADNDNTKAEAFNRIGPKSW--VKRNIAD	1059
Qy	64	AILGLDQKI-----SHLTDELDALFDDYOKARAVRRGLVEDNE-----	102
Db	1060	CFKLIRNKLTNQISDQPSHGNEUHELGHDEILADGLIKKGKEQTQLEVAIGDMFTI	1119
Qy	103	DSDSQSSFPVRYDKYNSGNTIDNFSLDTA-ITMPGTPRSDGDDALFFG-----DKSKQD	156
Db	1120	HGDMKNKPKKSKYLNATDD---DTASINYSYGSHKNRPKDESHKSGAETMEGEKRD	1175
Qy	157	ASNVDV---BELRQQAQWEE	174
Db	1176	ASKEDGLDDEEL-DEEGEECE	1195

RESULT 9
US-08-808-793-23
Sequence 23, Application US/08808793
Patent No. 5858713
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,793
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,361
FILING DATE: 24-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,649
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brame, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 2100 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-808-793-23

Query Match          9.1%  Score 85.5;  DB 2;  Length 2100;
Best Local Similarity 21.9%  Pred. No. 2.6;
Matches 44;  Conservative 36;  Mismatches 70;  Indels 51;  Gaps 9;

QY  4  SLRHVSTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREAS 63
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1029 ALLSNFGSSLSAPT-----ADNTNKIAEAFNRIGRFSW--VKRNAD 1072

QY  64  AILGLDQKI-----SHLTDDLDALFDVQKARAVRGLVEONE----- 102
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1073 CFKLIRNKLNTQISDPSEHGDNELGHEILADGLIKKIQEQTQLEVAIGDGMEFTI 1132

QY  103 DSDSQSPFVRDXYMSNCTIDNFSIDTA-ITMPTGTPRSDDGDALFFG-----DKSKQD 156
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1133 HGDKNKPKKSKYLNNATDD----DTASINSYSGHKNRPFKDESHKGSAAETMEGEKRD 1188

QY  157 ASNVDV---EELRQQAQME 174
    || || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  1189 ASKEDGLDEEL-DEEGECE 1208

RESULT 10
US-08-772-512A-19
; Sequence 19, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; TITLE OF INVENTION: FLIES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,618
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Braham, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/601(CRFD-1657)
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2100 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
```

```
;
; MOLECULE TYPE: protein
US-08-772-512A-19

Query Match          9.1%  Score 85.5;  DB 3;  Length 2100;
Best Local Similarity 21.9%  Pred. No. 2.6;
Matches 44;  Conservative 36;  Mismatches 70;  Indels 51;  Gaps 9;

QY  4  SLRHVSTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREAS 63
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1029 ALLSNFGSSLSAPT-----ADNTNKIAEAFNRIGRFSW--VKRNAD 1072

QY  64  AILGLDQKI-----SHLTDDLDALFDVQKARAVRGLVEONE----- 102
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1073 CFKLIRNKLNTQISDPSEHGDNELGHEILADGLIKKIQEQTQLEVAIGDGMEFTI 1132

QY  103 DSDSQSPFVRDXYMSNCTIDNFSIDTA-ITMPTGTPRSDDGDALFFG-----DKSKQD 156
    :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| : :|| :
Db  1133 HGDKNKPKKSKYLNNATDD----DTASINSYSGHKNRPFKDESHKGSAAETMEGEKRD 1188

QY  157 ASNVDV---EELRQQAQME 174
    || || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  1189 ASKEDGLDEEL-DEEGECE 1208

RESULT 11
US-08-264-002-5
; Sequence 5, Application US/08264002
; Patent No. 5559019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: CEHK
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1087
; US-08-264-002-5

Query Match          8.5%  Score 80;  DB 1;  Length 1087;
```

Best Local Similarity 22.8%; Pred. No. 4;
Matches 41; Conservative 27; Mismatches 54; Indels 58; Gaps 10;

QY 2 EQSLRHVNSVQSGPAGRHDTLSKAGSDADGDSRSDPLPHFRGKSDWYFARREA 61
Db 276 EQSL-----QSDGSCSDSDMNMSCVSADEEDVESQDDSD---FHVND-----ATEES 320
QY 62 ASATILGLDOKTS-----HLTD--ELDALFDDVQKARAVRGLVEDNEDSDSQS 108
Db 321 IDVSYSIESQAEASATEDLASCHNSNDKNEKVLVD-----EDTSKYD 364
QY 109 SFPPV-----RDKVMNSGNTDNFSLDTAITMPGTPRSDDDGDLFFGDKSKQDASNVVDVE 163
Db 365 NLPVEMSAKEGESEGTIDS-SVSSSTSSST---GDGD-----DSATSYDSEDIQ 415

RESULT 12

US-08-557-122A-26
; Sequence 26, Application US/08557122A
; Patent No. 5875664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-557-122A-26

Query Match 8.3%; Score 78; DB 2; Length 3052;

Best Local Similarity 24.0%; Pred. No. 30;
Matches 42; Conservative 26; Mismatches 77; Indels 30; Gaps 8;

QY 24 DRTLKAGSDADGDSRSDPLPHFRGKSDWYFARRAASAILGLDKIKSHL----- 75
Db 2874 DLQDGAGDDDDLEDEAEEDPLE-EDDDQKAVKDLQDG-AGDDDDLEDEAEEDPDM 2931
QY 76 -TDDELDAFDVQKARAVRGLVE-----DNEDSDSSSPVPRDKYNSNGTIDNFSL 127
Db 2932 EEDDDQKAVKDELDQDGAGDEGLDELEAEEDPLEDDDDQA--VRDELQDGAADD-DL 2988
QY 128 DTAITMPGTPRSDDDGDLFFGDKSKQDASNVVDVEELRQQAOMEALTKAQE 182
Db 2989 EDLETDEETLEGGDD-----EQIKQKDELHFDVD---GKALYEAEQAEE 3034

RESULT 13

US-08-531-525-33
; Sequence 33, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discopyge ommata
US-08-531-525-33

Query Match 8.2%; Score 76.5; DB 2; Length 200;

Best Local Similarity 23.1%; Pred. No. 0.84;
Matches 28; Conservative 21; Mismatches 49; Indels 23; Gaps 5;

QY 81 DALFDDVQ--KARAVRGLVEDNE-----DSDSQSPFPVRDKYNSNG-----TIDNF 125
Db 35 DEFVEDYPTKADSYRKVKVLDGEVQIDILDTAGOEDYAINDNVFRSGEGFLCVFSIEQ 94
QY 126 SLDTAITMPG---TPRSDDDGDLFFGDKSKQDASNVVDVEELRQQAQ-----MEELK 177
Db 95 SFTATVFRQILRVKEEDKIPLLVGNKSDLEDRRQVSIIEAKSAEKGQVIVETSAK 154
QY 178 T 178
Db 155 T 155

RESULT 14

US-08-718-270A-33
; Sequence 33, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:14:12 ; Search time 46.78 Seconds

(without alignments)
268.839 Million cell updates/sec

Title: US-09-235-416-1_COPY_602_784

Perfect score: 938

Sequence: 1 OBQSULRHVSTNSQLGPAP.....ELRQQQAQMEALKTAQKEF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_67.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	100	10.7	2649	2 T51023	hypothetical prote
2	98.5	10.5	1883	2 T13944	chromodomain-helic
3	93.5	10.0	599	2 T39990	hypothetical prote
4	93.5	10.0	1254	2 T41262	mutS family DNA mi
5	91.5	9.8	3839	2 T49799	related to TOM1 pr
6	90.5	9.6	1707	2 T18951	hypothetical prote
7	90	9.6	611	2 T06458	nucleolin homolog
8	89.5	9.5	590	2 A56535	gadd34 protein - l
9	89.5	9.5	1070	2 T30848	Duffy receptor - p
10	89	9.5	476	2 T36395	probable transcrip
11	88.5	9.4	1353	2 T00347	hypothetical prote
12	88	9.4	384	2 S31796	vasodilator-stimul
13	87.5	9.3	233	2 A36425	clathrin light cha
14	87.5	9.3	933	2 S41539	fibrinogen-binding
15	87	9.3	1302	1 JC6009	surface-located me
16	86.5	9.2	693	2 T01122	hypothetical prote
17	86	9.2	1957	2 T38077	hypothetical coile
18	85.5	9.1	438	2 B39487	T-complex protein
19	85.5	9.1	629	2 B64075	transcription init
20	85.5	9.1	805	2 T06657	hypothetical prote
21	85.5	9.1	1820	2 A33299	sodium channel pro
22	85.5	9.1	2946	2 T00867	hypothetical prote
23	85	9.1	747	2 T33488	hypothetical prote
24	84	9.0	665	3 JC7191	85K c-Cbl-interact
25	83.5	8.9	415	2 S35760	fcrA protein precu
26	83.5	8.9	438	2 S13637	hypothetical prote
27	83.5	8.9	1203	2 B55094	chromosomal protei
28	83.5	8.9	1304	2 T19397	hypothetical prote
29	83	8.8	732	2 T32757	hypothetical prote

30	83	8.8	1051	2 T18302	apsB protein - Eme
31	83	8.8	1757	2 T05204	hypothetical prote
32	82.5	8.8	418	2 S28410	t-complex protein
33	82.5	8.8	533	2 T06153	hypothetical prote
34	82	8.7	374	2 T33328	hypothetical prote
35	82	8.7	598	2 T11615	hypothetical prote
36	82	8.7	1130	2 T34081	hypothetical prote
37	82	8.7	1136	2 A56559	enhancer-trap-locu
38	81.5	8.7	628	2 S61160	hypothetical prote
39	81.5	8.7	655	2 S18349	dnak-type molecula
40	81.5	8.7	666	2 S34197	DNA-binding protei
41	81.5	8.7	809	2 S32899	ferric-pseudobacti
42	81.5	8.7	852	2 A28459	cell surface glyco
43	81.5	8.7	1063	2 A40253	acidic nuclear pro
44	81	8.6	635	2 T09648	nucleolin homolog
45	81	8.6	772	2 S22660	excision repair pr

ALIGNMENTS

RESULT 1

T51023

hypothetical protein B7F21.40 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000

C:Accession: T51023

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu submitted to the Protein Sequence Database, July 2000

A:Reference number: Z52886

A:Status: preliminary

A:Accession: T51023

A:Molecule type: DNA

A:Residues: 1-2649 <SCH>

A:Cross-references: EMBL:AL389901; GSPDB:GN00116; NCSP:B7F21.40

A:Experimental source: BAC clone B7F21; strain OR74A

C:Genetics:

A:Gene: NCSP:B7F21.40

A:Map position: 6

A:Introns: 1619/3; 2584/1

Query Match	10.7%	Score 100;	DB 2;	Length 2649;
Best Local Similarity	25.1%	Pred. No. 6.5;		
Matches	54;	Conservative	29;	Mismatches 80; Indels 52; Gaps 10;
QY	10	VTNSQLGSPAG	-----RHDTLSKAGSD-----ADGDSRS-----DSP	43
Db	559	VQSQHGRRPFGGHRRPSSSHDHGGETERRRSPDPAKSESHLVAAQSRSSLLTGANDTP	618	
QY	44	LPHPFRGKDSMFYARREASAILGLDQKI-----SHLTDDLDALFDVQKARVRGLV	98	
Db	619	IRPERCQTQASARASVORDTRASLDNRDNRDTSFGGAHRSFELDA---ERRDERPALMAPM	675	
QY	99	EDNEDSSS-----OSSFPVRDK--YASNGTIDNFSLDTAIT--MPCPTPSDDDDGDLFGCD-	150	
Db	676	ROREDADKSRIPESRVPVKRRSRVLAIPSTRLQLPTQOTSLPQT-ESDDDDMDDDYFDA	734	
QY	151	-----KKSKQDASNVYVEELRQQQAQMEAL	176	
Db	735	EITTKQETELKAKLADSTAGVPMQIVRYACTVHDAM	769	
RESULT 2				
T13944				
Chromodomain-helicase-DNA-binding protein - fruit fly (Drosophila melanogaster)				
C:Species: Drosophila melanogaster				
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000				
C:Accession: T13944				
R:Stokes, D.G.; Tartof, K.D.; Perry, R.P.				
Proc. Natl. Acad. Sci. U.S.A. 93, 7137-7142, 1996				
A:Title: CHD1 is concentrated in interbands and puffed regions of Drosophila polytene				
A:Reference number: Z17823; MUID:96293489				

A:Title: Cloning of the Plasmodium vivax Duffy receptor.

A:Reference number: Z20901; MUID:91187056

A:Accession: T30848

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1070 <FAN>

A:Cross-references: EMBL:M61095; NID:g160275; PID:g160276; PIDN:AAA63423.1

C:Genetics:

A:Introns: 19/3; 1006/1; 1032/2; 1057/1

A:Note: PVDR

Query Match 9.5%; Score 89.5; DB 2; Length 1070;
Best Local Similarity 24.0%; Pred. No. 16;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

QY 6 LRHSVTSNQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREASAI 65

Db 671 LRHSKNSDSGPA-----ESMANPDSNKGTEG-----KGQNDMAKATKSSNS 717

QY 66 LGLDQKTSHTLTDDELDFDVQKARVRGLVEDNEDSDSQSFPPVRDKYMSN---GTI 122

Db 718 DG-----TSSATGTTDAV-----DREINKGVPE-----RDKTVGSKDGGGE 755

QY 123 DNFSLDTAITMPGTPRSDDDGDALFFGDKKSKQDASNDVVEE---LRQQAQMEAEALKTA 179

Db 756 DNSANKDAATVVGEDRLRENSAG-----GSTNDRSKNDTEKNGASTPDSKQSDATAIS 809

QY 180 KQE 182

Db 810 KTE 812

RESULT 10

T36395

probable transcription regulator - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999

C:Accession: T36395

R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21605

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-476 <MUR>

A:Cross-references: EMBL:AL109987; PIDN:CAB53399.1; GSPOB:GN00070; SCOEDB:SCF20.01c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCF20.01c

Query Match 9.5%; Score 89; DB 2; Length 476;
Best Local Similarity 36.2%; Pred. No. 6.5;
Matches 21; Conservative 7; Mismatches 30; Indels 0; Gaps 0;

QY 5 LLRHSVTSNQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREAA 62

Db 268 LLELAFTDRPADAPGPGTGSRTGTRHGSRAEGGAAGRPPLGLAGTDPANWSVSRTAA 325

RESULT 11

T00347

hypothetical protein DKFZp566G1246.1, version 1 - human (fragment)

N:Alternate names: hypothetical protein KIAA0592

C:Species: Homo sapiens (man)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000

C:Accession: T00347

R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A:Title: Prediction of the coding sequences of unidentified human genes. IX. The completed

A:Reference number: T14086; MUID:98290545

A:Accession: T00347

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1353 <NAG>

A:Cross-references: EMBL:AB011164; NID:g3043707; PIDN:BAA25518.1; PID:g3043708

A:Experimental source: brain; clone HU2807

C:Genetics:

A:Note: KIAA0592

Query Match 9.4%; Score 88.5; DB 2; Length 1353;
Best Local Similarity 24.4%; Pred. No. 25;
Matches 50; Conservative 19; Mismatches 61; Indels 75; Gaps 9;

QY 13 SQSGSPAPGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREASAILGLDOKI 72

Db 474 SPYGPPPTGLFD-----DDGDD-----DDDFSAPHSKPSKTRKV-QST 512

QY 73 SHLTDDDELDALE-----DDVQKARA-----VRRGLIVE 99

Db 513 ADIFGDEGDLFREKAVASPEATVSQTDENKARAEKKVTLSSSKNLKPSSETKTQKGLFS 572

QY 100 DNEDSD-----SQSFPVRDKYMSNGTTDNFSLDTAITMPGTPRSDDDGDALFFGDKKSKQ 155

Db 573 DEEDSEDLFSQSASNLKGALLPG-----KLPTSVSLFD-----DEDEDNLFGGTAAKKQ 624

QY 156 DASNVDEELRQQAQMEAEALKTA 180

Db 625 TLS-----LQAOREEKAKASE 640

RESULT 12

S51796

vasodilator-stimulated phosphoprotein VASP - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C:Accession: S51796

R:Haffner, C.; Jarchau, T.; Reinhard, M.; Hoppe, J.; Lohmann, S.M.; Walter, U.

EMBO J. 14, 19-27, 1995

A:Title: Molecular cloning, structural analysis and functional expression of the prol

A:Reference number: S51796; MUID:95129547

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-384 <HAF>

A:Cross-references: EMBL:Z46388; NID:g624249; PIDN:CAA86522.1; PID:g624250

C:Keywords: phosphoprotein

Query Match 9.4%; Score 88; DB 2; Length 384;
Best Local Similarity 25.6%; Pred. No. 6.1;
Matches 52; Conservative 26; Mismatches 75; Indels 50; Gaps 9;

QY 18 PAFGRHRTLSKAGSDAGDSRSDSPLPFRGKDSDFYARREASAILGLDOKISHLTD 77

Db 187 PPGVSLSGSAGAGGAGGPPAPPLPTAGTSGGTGAPGLAA-AIAG-AKLKRVSK 243

QY 78 DELDALFDDVQKARVRR---GLVED-----NEEDSDSQSF 110

Db 244 QEEASGPPVPKAESTRSTGGGLMEENAMLRARRKATQVGEKPPKDESANEPEARVPV 303

QY 111 P-----VRDKYMSNGTI-----DNFSLDTAITMPGTPRSDDDGDALFFGDKKSKQDASNV 160

Db 304 PAQSETVRRPWEKNSTPLPRMKSSSVTTSEAHPTSPSSDESDEL-----ERVQEL--- 355

QY 161 DVEELRQQ-QAQMEEALKTAQE 182

Db 356 -LEEVRKELQKVKEIEIAFVQE 377

RESULT 13

A36425

clathrin light chain - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein G7056; protein YGR167w

Search completed: April 25, 2001, 10:14:18
Job time: 294 sec

```

F.25-1302/Product: surface-located membrane protein Lmp3 #status predicted <MAT>
F.957-992/Domain: tetratricopeptide repeat homology <TT1>
F.993-1026/Domain: tetratricopeptide repeat homology <TT2>
F.1089-1120/Domain: tetratricopeptide repeat homology <TT3>
F.1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match          9.3%;   Score 87;   DB 1;   Length 1302;
Best Local Similarity 20.9%;   Pred. No. 32;
Matches 44;   Conservative 40;   Mismatches 83;   Indels 44;   Gaps 8;

QY  7 RHVSYNSQLGSPAPGRHRTL-----SKAGSDAGDSRS-----DSPLPHFR 48
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  444 KKSVTNSSNKSDSIITAAEALIQALDANKAKQDVDEANKSIKEQNALDKANTLLPQLN 503
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  49 GKQSDWFYARREASATILGDQIKISHLTDDELDAL---FDDVQKARAVRGILVENEDSD 105
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  504 DNDSEIVAKESLNAEITNANKAVQNQDNASQSAKSSILD--KVTKIQNLQTEFNKDKD 561
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  106 SQSPFPVRKYNSNGTIDNFSLDITATWPG-----TPRSDDDGDALFFGDKKSKODAS 158
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  562 AK----FKELEGRKIDINPLFDDVYNNPNVATLVKLTNKAODDKSVTKSNKNSKIIAA 617
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  159 NVDVEELRQ-----QQAQMEAEALTKAQE 182
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  618 N---DELQALDKAKVAKQIDDEANKSKTKQ 645
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

FT MOD_RES 153 153 PHOSPHORYLATION (BY CAPK AND CGPK) (BY
FT FT SIMILARITY).
FT MOD_RES 235 235 PHOSPHORYLATION (BY CAPK AND CGPK) (BY
FT FT SIMILARITY).
FT MOD_RES 274 274 PHOSPHORYLATION (BY CAPK AND CGPK) (BY
FT FT SIMILARITY).
FT CONFLICT 209 209 T -> A (IN REF. 2).
FT CONFLICT 288 288 MISSING (IN REF. 2).
SQ SEQUENCE 376 AA; 39824 MW; CC338D07519A0294 CRC64;

Query Match 11.4%; Score 106.5; DB 1; Length 376;
Best Local Similarity 26.0%; Pred. No. 0.12;
Matches 52; Conservative 28; Mismatches 75; Indels 45; Gaps 8;

QY 18 PAPGRHRTLSKAGSDAGDSRSDSPLPHPF--RGKSDWYARREASAILGLDQKISHLT 76
Db 180 PPPGLPSSGVSGAGHAGAAPPPAPPLPTTQGPNSGGSGAPGLAA-AIAGAKLRKVKSQE 238
QY 77 DDELDAFDVOKARAVRRGLVED-----NEDSDSQ---S 108
Db 239 EASGGPLAPKAENSRSTGGGLMEEMNAMLARRKATQVGEKPKDESASGESEARLPAQ 298
QY 109 SFPRYDKYMSNGTT-----DNFSLDTAITMPGTPRSDDGDGALFFGDKKQDASNVDFVE 163
Db 299 SEPVRPWEKNSITLPRMKSSSVTTSEAHPTPCSSDDSDL-----ERVQEL----LE 349
QY 164 ELROQ-QAQMEEALKTAQKE 182
Db 350 EVRKELOMKKEIEIEVFQKE 369

RESULT 2
IE63_MCMVS STANDARD; PRT; 841 AA.
ID IE63_MCMVS
AC Q69154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE TRANSCRIPTIONAL REGULATOR IE63 HOMOLOG (PROTEIN UL69).
GN UL69.
OS Murine cytomegalovirus (strain Smith).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10367;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96082764; PubMed=7483291;
RA Messerle M., Rapp M., Lucin P., Koszinowski U.H.;
RT "Characterization of a conserved gene block in the murine
cytomegalovirus genome.";
RL Virus Genes 10:73-80(1995).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HVS-1 57, HCMV AND MCMV
CC UL69.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L07319; AAA9666.1;
KW Transcription regulation.
SQ SEQUENCE 841 AA; 92706 MW; 9F18372739D723E4 CRC64;

Query Match 9.8%; Score 91.5; DB 1; Length 841;
Best Local Similarity 25.9%; Pred. No. 5;
Matches 43; Conservative 15; Mismatches 41; Indels 67; Gaps 7;

```

```

QY 17 SPAPGRHRTLSKAGSDAGDSRSDSPLPHPF--RGKSDWYARREASAILGLDQKISH 74
Db 622 SPSPASRRRPSK-----SPSSPRPHDPSPGEPAD---AKEELATA----- 660
QY 75 LTDDLDALFDVOKARAVRRGLVEDNEDSDSQSFVPRYDKYMSNGTIDNFSLTAITMP 134
Db 661 --GDEGEVRSRPGCSVATRRGSAD--ESSDSSS----- 692
QY 135 GTPRSDDGDGALFFGDKKQDASNVDFVELRQ---QAQMEELAK 177
Db 693 -----SSDSSSSSDEESDVEDCRELDQSKRLEALE 725

RESULT 3
PVDR_PLAYS STANDARD; PRT; 1070 AA.
ID PVDR_PLAYS
AC P22290;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DUFFY RECEPTOR PRECURSOR (ERYTHROCYTE BINDING PROTEIN).
GN PVDR.
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5856;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91187056; PubMed=1849231;
RA Fang X., Kaslow D.C., Adams J.H., Miller L.H.;
RT "Cloning of the Plasmodium vivax Duffy receptor.";
RL Mol. Biochem. Parasitol. 44:125-132(1991).
CC -1- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
CC DETERMINANT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: HIGH, TO P.KNOWLEDGE DUFFY RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M61095; AAA63423.1;
KW Malaria; Receptor; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1070 DUFFY RECEPTOR.
FT DOMAIN 21 1007 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1008 1025 POTENTIAL.
FT DOMAIN 1026 1070 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 715 715 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 787 787 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 825 825 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 903 903 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1070 AA; 119683 MW; CB051DF13E294603 CRC64;

Query Match 9.5%; Score 89.5; DB 1; Length 1070;
Best Local Similarity 24.0%; Pred. No. 9.6;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

QY 6 LRISVTNQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPHPFGRKSDWYARREASAI 65
Db 671 LRHSKDNDSQSDGPA-----ESMANPDSNKGTEG---KQDNDMAKATKSSNSS 717
QY 66 LGLDQKISHLTDELDALFDVOKARAVRRGLVEDNEDSDSQSFVPRYDKYMSN---GTI 122

```


Db 26 QGVLTFPAEINLLPDAIDPEYDKLLOTLQNDAGIPVLDEAPESDDMLSDTIPDED-- 83
Qy 56 YARREAAAILGLDQKISHLTDD-----ELDALFDVQKARAVRGLVEDNES 104
Db 84 -AVEEATOILSNVESEIGRTDPVRMYRMGVTLTREDSEISAKRIEGGI---DEVQ 139
Qy 105 DSQSFVRDKYMSNGTIDNF-----SLDTAITMPTGPRSDDDGDALFFGDKKS 153
Db 140 TSISAYP-----EALNGLLKNYDDVEKGNFRILTDLITGFVDPNAAIEEHNGLDFDSEDD 195
Qy 154 QDASNVDEELRQOQAQMEEA 175
Db 196 ERESSNADVEDNEDENESES 217
RESULT 8
XCPE_XENLA STANDARD; PRT; 1203 AA.
ID XCPE_XENLA AC P50533;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CHROMOSOME ASSEMBLY PROTEIN XCAP-E.
GN XCAP-E.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95042742; PubMed=7954811;
RA Hirano T., Mitchison T.J.;
RT "A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro."
RL Cell 79:449-458(1994).
CC -1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF MITOTIC CHROMOSOMES.
CC -1- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR EMBL: U13674; AAA64680.1; --
DR HSSP: P07751; 1A33
KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.
FT NF-BIND 32 39
FT DOMAIN 172 356
FT DOMAIN 397 513
FT DOMAIN 670 1032
FT DOMAIN 1095 1121
FT SEQUENCE 1203 AA; 136341 MW; 043223DD0027DF309 CRC64;
Query Match 8.9%; Score 83.5; DB 1; Length 1203;
Best Local Similarity 20.3%; Pred. No. 33;
Matches 29; Conservative 29; Mismatches 66; Indels 17; Gaps 3;
Qy 47 FRGKDSWFYARREAAAILGLDQKISHLTDDDELDAFDVQKARAVRGLVEDNESDS 106
Db 713 YRQLQKQWEMKSEAEILLQTKLQSSYHKQOEELSLKQTIIESEETLKNTKEVQKAAE 772

Qy 107 OSSFPVRDKYMSNGTID-----NFSLDTAITMPTGTPRSDDDGDALFFGDKKSQDAS 158
Db 773 K-FKVLHKKMNAERERELKEAQKLDTA-----KKKADASNKKKEKQOEVDAL 823
Qy 159 NVDVBEELRQOQAQMEALKTAQ 181
Db 824 VLELEELKREQTTTKQIETVDE 846
RESULT 9
SNWA_DICDI STANDARD; PRT; 685 AA.
ID SNWA_DICDI AC P54705;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SNWA PROTEIN.
GN SNWA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97128797; PubMed=8973337;
RA Folk P., Puta F., Krpejsova L., Blahuskova A., Markos A.,
RA Rabino M., Dottin R.P.;
RT "The homolog of chromatin binding protein Bx42 identified in Dictyostelium."
RL Gene 181:229-231(1996).
CC -1- SIMILARITY: BELONGS TO THE SNW FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR EMBL: U43887; AAB40497.1; --
DR Dictyob; DD00074; snwa.
FT DOMAIN 31 41
FT DOMAIN 194 360
FT DOMAIN 245 253
FT DOMAIN 409 415
FT DOMAIN 539 616
FT SEQUENCE 685 AA; 78529 MW; 1DC8521E9997A583 CRC64;
Query Match 8.8%; Score 82.5; DB 1; Length 685;
Best Local Similarity 24.9%; Pred. No. 20;
Matches 48; Conservative 28; Mismatches 86; Indels 31; Gaps 8;
Qy 1 QEOILR---HSVTSQSGSPAPGRHDTLSKAGSDAGDSRSDSPLPHFRGKSDWFYA 57
Db 339 RKQMLRKLAEADVNRSGI-----IQQRTRKDSDSDNDNDSSDSDKNTPPMNR 394
Qy 58 RREAASAILGLDQKISHLTDDDELDAFDVQKARAVRGLVEDNES-----DSQSF 110
Db 395 RSRSRST-----ERIPSRNDNDD--DDRYIKDNRDNRDNRDNRDNRDNRDNR 446
Qy 111 PVRKYMSNGTIDN-FSLDTAITMPTG--TPRSDDDGDALFF-----GDKKSQDASNV 161
Db 447 DSRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSDRSD 506
Qy 162 VEELRQOQAQME 174
Db 507 DERVRERKEKLE 519
RESULT 10
YAA5_SCHPO

```

ID YAA5_SCHPO STANDARD; PRT; 598 AA.
AC Q09759;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHEICAL 69.5 KDA PROTEIN C22G7.05 IN CHROMOSOME 1.
GN SPAC22G7.05.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA badcock K.; Churher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO YEAST N0388.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z54328; CAA91129.1; -.
KW Hypothetical protein.
FT DOMAIN 104 107 POLY-GLU.
FT DOMAIN 243 246 POLY-GLU.
FT SEQUENCE 598 AA; 69523 MW; 5C6A68D67AC0FB5 CRC64;
SQ
Query Match 8.7%; Score 82; DB 1; Length 598;
Best Local Similarity 22.9%; Pred. No. 19;
Matches 44; Conservative 18; Mismatches 66; Indels 64; Gaps 6;
QY 31 GSDADGSRSDSPHFRGKDSDFYARRRASAAILGDKISHLTDDELDLDA---LFDDV 87
DB 92 GVDGEGDESSE-----EEDSDGELVTPVDAAILRMIVKIRNKDPDLYDSQKYPDEV 147
QY 88 QKARAVRRGLVEDNEDSDSSFPVRDKYMS-----NGTIDNFSLDTAITMPG 135
DB 148 EK-----DVGSLKSKDGFRSVLYKDYHRKLLSGEILDAEEDPWPND 192
QY 136 TP-----RSDGDDALFFGDKKSKQ-DASNVVDVEEL 165
DB 193 NPTHVERQERLRKETIAAFHDVNGNKDAVNSDEDDGDLVKKETKKQLEEEHGYVERF 252
QY 166 RQQAQMEELK 177
DB 253 LLESQAQSKARK 264
RESULT 11
Z151_CHICK STANDARD; PRT; 706 AA.
AC Q90625;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC FINGER PROTEIN 151 (ZINC FINGER PROTEIN 213) (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96003919; PubMed=7575457;
RA Schulz T.C., Hopwood B., Rathjen P.D., Wells J.R.E.;
RT "An unusual arrangement of 13 zinc fingers in the vertebrate gene
```

```

RT 213 ";
RL Biochem. J. 311:219-224(1995).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN WHICH MODULATES DNA-BINDING.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14555; AAA21556.1; -.
DR HSSP; P15822; 1BBQ.
DR InterPro; IPR000822; -.
DR Pfam; PF00096; zf-C2H2; 13.
DR PROSITE; PS50097; BTB; PARTIAL.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 13.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 12 BTB.
FT DOMAIN 205 640 ZINC_FINGERS.
FT ZN_FING 205 227 C2H2-TYPE.
FT ZN_FING 233 255 C2H2-TYPE.
FT ZN_FING 261 283 C2H2-TYPE.
FT ZN_FING 289 311 C2H2-TYPE.
FT ZN_FING 317 339 C2H2-TYPE.
FT ZN_FING 345 367 C2H2-TYPE.
FT ZN_FING 373 395 C2H2-TYPE.
FT ZN_FING 401 423 C2H2-TYPE.
FT ZN_FING 427 450 C2H2-TYPE.
FT ZN_FING 457 479 C2H2-TYPE.
FT ZN_FING 485 507 C2H2-TYPE.
FT ZN_FING 513 536 C2H2-TYPE.
FT ZN_FING 618 640 C2H2-TYPE.
SQ SEQUENCE 706 AA; 76820 MW; 4BDAAC68808F1136 CRC64;
Query Match 8.7%; Score 82; DB 1; Length 706;
Best Local Similarity 24.4%; Pred. No. 23;
Matches 29; Conservative 22; Mismatches 46; Indels 22; Gaps 4;
QY 10 VTNSQLGSPA---PGRHRTLSKAGSDADGSRSDSPHFRGKDSDFYARRRASAAIL 66
DB 87 VDNAEASPAVSPSPQPAEVEVGNSSPGKSGDAPSTEARGNELE----- 133
QY 67 GLDQKISHLTDDELDLDFDDVQKARAVRRGLVEDNE-----DSDSQSFPVRDKYMSNGT 121
DB 134 GKEEGEAMVEDEEEAKIPKAAQPKSKENAEADNESGSTDSGQENSGETR--LLRSGT 190
RESULT 12
YL17_CAEEL STANDARD; PRT; 1130 AA.
AC Q11102;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHEICAL 131.5 KDA PROTEIN C02F12.7 IN CHROMOSOME X.
GN C02F12.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Miller N.;
```


Mon Apr 30 07:47:12 2001

us-09-235-416-1_copy_602_784.rsp

Page 9

Db 549 ALENY-----AYNMRNTIKDDKIPGKLDAGDKKKTETAVNEAIEWLEKNQLAEVDELEDKL 604

Qy 177 K 177

Db 605 K 605

Search completed: April 25, 2001, 10:18:38
Job time: 304 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: April 25, 2001, 10:18:12 ; Search time 77.5 Seconds
(without alignments)
276,762 Million cell updates/sec

Title: US-09-235-416-1_copy_602_784
Perfect score: 938
Sequence: 1 QGSLLRHSYTNQLSPAP.....ELRQQQOMEEALTKAQEF 183

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues.

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL15.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_oranella.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_unclassified.*
- 13: sp_vertebrate.*
- 14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	11.5	375	11 Q9R214	Q9r214 mus musculu
2	106.5	11.4	376	11 P70460	P70460 mus musculu
3	100	10.7	2649	3 Q9P300	Q9p300 neurospora
4	98.5	10.5	1883	5 Q24376	Q24376 drosophila
5	98.5	10.5	1883	5 Q9VQ09	Q9vqj9 drosophila
6	93.5	10.0	599	3 Q74354	Q74354 schizosacch
7	93.5	10.0	972	5 Q9VBX1	Q9vbx1 drosophila
8	93.5	10.0	1254	3 Q74502	Q74502 schizosacch
9	91.5	9.8	3839	3 Q9P421	Q9p421 neurospora
10	91	9.7	219	4 Q9N223	Q9n223 homo sapien
11	90.5	9.6	1707	5 Q17652	Q17652 caenorhabdi
12	90	9.6	611	10 Q41042	Q41042 plasm mod
13	90	9.6	757	5 Q26189	Q26189 plasm mod
14	89.5	9.5	590	11 Q60465	Q60465 cricetus
15	89.5	9.5	750	5 Q26188	Q26188 plasm mod
16	89.5	9.5	788	10 Q9X183	Q9x183 arabidopsis
17	89	9.5	476	2 Q9S2E5	Q9s2e5 streptomyce
18	89	9.5	757	5 Q26127	Q26127 plasm mod
19	89	9.5	757	5 Q26186	Q26186 plasm mod

20	89	9.5	758	5 Q26125	Q26125 plasm mod
21	89	9.5	759	5 Q26128	Q26128 plasm mod
22	88.5	9.4	923	4 Q9NVB8	Q9nvb8 homo sapien
23	88.5	9.4	1353	4 Q9YAE1	Q9yae1 homo sapien
24	88	9.4	580	5 P91156	P91156 caenorhabdi
25	88	9.4	592	11 Q9JL09	Q9jlu9 rattus norv
26	88	9.4	631	5 Q9VAF4	Q9vaf4 drosophila
27	88	9.4	645	11 Q5JKQ0	Q5jkq0 rattus norv
28	88	9.4	1530	4 Q43241	Q43241 homo sapien
29	87.5	9.3	841	10 Q9LME4	Q9lmf4 arabidopsis
30	87.5	9.3	933	2 Q53653	Q53653 staphylococ
31	87	9.3	489	4 Q9NWH8	Q9nwh8 homo sapien
32	87	9.3	946	10 Q9M946	Q9m946 arabidopsis
33	87	9.3	1302	2 Q49547	Q49547 mycoplasma
34	86.5	9.2	302	10 Q9SUG6	Q9sug6 arabidopsis
35	86.5	9.2	693	10 Q50071	Q50071 arabidopsis
36	86.5	9.2	1539	4 Q9P2D2	Q9p2d2 homo sapien
37	85.5	9.1	437	5 Q9VNP3	Q9vnp3 drosophila
38	85.5	9.1	805	10 Q9TOH5	Q9toh5 arabidopsis
39	85.5	9.1	1030	10 Q9LHT5	Q9lht5 arabidopsis
40	85.5	9.1	2946	10 Q64634	Q64634 arabidopsis
41	85	9.1	747	5 Q9PZL0	Q9pzt0 caenorhabdi
42	85	9.1	758	5 Q26129	Q26129 plasm mod
43	85	9.1	1309	5 Q9VIV0	Q9viv0 drosophila
44	85	9.1	2186	5 Q9N906	Q9n906 trypanosoma
45	84	9.0	203	5 Q01721	Q01721 trichostrom

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	375 AA.
Q9R214	AC Q9R214			
ID	Q9R214			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VASODILATOR-STIMULATED PHOSPHOPROTEIN.			
GN	VASP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J;			
RA	Collins S.P., Uhler M.D.;			
RT	Cyclic AMP- and Cyclic GMP-dependent Protein Kinases Differ in Their			
RT	Regulation of Cyclic AMP Response Element-dependent Gene			
RL	J. Biol. Chem. 274:8391-8404(1999).			
DR	EMBL; AF084548; AAD16045.1;			
DR	MGD; MGI:109268; Vasp.			
DR	INTERPRO; IPR000697;			
DR	INTERPRO; IPR001960;			
DR	INTERPRO; IPR002965;			
DR	PFAM; PF00568; WH1; 1.			
DR	PRINTS; PR01217; PRICHTENS.			
SQ	SEQUENCE 375 AA; 39666 MW; 19369286CF4276C7 CRC64;			
Query Match	11.58;	Score 108;	DB 11;	Length 375;
Best Local Similarity	26.18;	Pred. No. 0.099;		
Matches	52;	Conservative	29;	Mismatches 74; Indels 44; Gaps 8;
QY	18	PAPGRDRTLKSGADGDSRSDSPHFRGKSDMFYARREASAILGLD-QKISHLT 76		
Db	180	PPPLPSPGSGAGHACACAPPAPPLPTAQGPSGGGAPGLAA-AIAGAKLRKVKQE 238		
QY	77	DELDALFDVQKAVRGRIVED-----NEEDSQ---SS 109		
Db	239	EASGGPLAKAENSNSKSTGGGLNEMNMLARRRKATQVGEKPKPKDESSEARLPAQS 298		

RESULT 3
Q9P3J0
ID Q9P3J0 PRELIMINARY; PRT; 2649 AA.
AC
CP Q9P3J0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN B7F21.40.
GN
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.

RESULT

Query Match	10.5%	Score 98.5	DB 5	Length 1883
Best Local Similarity	23.1%	Pred. NO. 4.4		
Matches 42	Conservative 24	Mismatches 73	Indels 43	Gaps 6
QY 8	HSYNTSQLGSPAPGRHRTILSKAGSDAGDSRS-----DSPLPHPFGKGDSDMFYARRE 60			
DB 29	HSGSGSGSGS-----SGSDSDSGSGNSDGRSSPEPEKSLSVAGFPPTAA 76			
QY 61	AASAILGLDOKISHLTDELDALEFDVQKARVRCLEVDEDNDSQSSFFPVROKYKNSG 120			
DB 77	AAQA-----DSTNGFTDQEDSSDG-----SSGSDSDSDAGSPDQR---NQ 117			
QY 121	TIDNFLDTATIMFTGPRSDGDGLFFGDKKSKQDASNVDEBELROQAQNEEALAKTAK 180			
DB 118	SINNANTSLLPKPEQNEEDNET-----EAGQQQAPASADASSDSSANVSPTSSSS 172			
QY 181	QE 182			
DB 173	SE 174			
RESULT 6				
074354				
AD 074354	PRELIMINARY;	PRT;	599 AA.	
IC 074354				
DT 01-NOV-1998	(TRENBLrel. 08, Created)			
DT 01-NOV-1998	(TRENBLrel. 08, Last sequence update)			
DT 01-NOV-1998	(TRENBLrel. 08, Last annotation update)			
DE	HYPOTHETICAL 66.0 KDA PROTEIN.			
GN	SPBC25D12.02C.			
OC	Schizosaccharomyces pombe (Fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;			
OC	Schizosaccharomycetaceae; Schizosaccharomycetes.			
OC	NCBI_TaxID=4896;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=972H-;			
RA	Lyme M., Rajandream M.A., Barrell B.G., Whithead S., Chillingworth T., Church C.M.			

```

RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AL031158; CAA20098.1; -.
KW Hypothetical protein.
SQ SEQUENCE 599 AA; 65977 MW; 6BF1P3AC1FC689E8 CRC64;

Query Match 10.0%; Score 93.5; DB 3; Length 599;
Best Local Similarity 19.8%; Pred. No. 3.1;
Matches 52; Conservative 29; Mismatches 72; Indels 109; Gaps

QY 8 HSVNLSQSPAPGRHRTLSKAGSDAGDSRSDSPIL-----PFRGKSDMFYARRE 61
DQ || || || || || || || || || || || || || || || || || || ||
Db 217 HOSVNSLTSPQPTTYNANFEFSINDASSDSDAPLTLSSPSRLMKDNRKYLVEHS 276
QY 62 ASAIL-----GLDQK----- 71
DQ || || || || || || || || || || || || || || || || || || ||
Db 277 PAALIKESITDGDIDKSLRSSTREVSVEPNEDSVNDSSDVSDEKETAEKHEIRAPA 336
QY 72 -----ISH-----LTDDELAL-----FDDVQKARVRRL 97
DQ || || || || || || || || || || || || || || || || || || ||
Db 337 IIVRETSSHSSTAVPSNDFTSENDTLSSSTTSISSPSSENSDTSDDLTKVDSPNKSL 396
QY 98 VEDN-----EVDSSGSPFVRDKYMSNGTIDNFSLDTAITMPTSPRSDGDLFFGD 150
DQ || || || || || || || || || || || || || || || || || || ||
Db 397 VDNVSAKHDKSENGSKSPPPSOTLV-----TSTISAAGNEPSDEIGSE---ND 445
QY 151 KSKQDA-SNVDVEELRQQAQ 171
DQ || || || || || || || || || || || || || || || || || || ||
Db 446 SDSDSDDSSVPLSQIQKKSQQ 467

RESULT 7
Q9VBX1 ID Q9VBX1 PRELIMINARY; PRT; 972 AA.
AC Q9VBX1
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CG11847 PROTEIN.
GN CG11847.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Brokstein P., Brothier P.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```

SQ SEQUENCE 1254 AA; 141510 MW; 7A28EC8DD8D2845B CRC64;

Query Match 10.0%; Score 93.5; DB 3; Length 1254;
Best Local Similarity 20.5%; Pred. No. 7.4;
Matches 38; Conservative 33; Mismatches 71; Indels 43; Gaps 5;

QY 12 NSQLGSPAGRHRTLSKAGSDADGSDSPPLPHRGRKDSWTFYA----RREAASAILG 67
DB 121 HDSLGEESPGKLLRTSVKQEPDSEEDISPTKKKSPKSLDTSIFQAEDQFRHPVSKLEN 180
QY 68 LD-----QKISHLTDELDFDDVOKARAVRGLVEDNEDSD----- 105
DB 181 SELSEVDKFTASRRKKPVSAESDE-DEDFODAPTGRSHKRIVSDSDSDYVEPDHI 239
QY 106 ----SQSFFPVROKYMNSGTIDNFSLOTATIMPTGTPRSDDDGDALFFGDKSKSQASNV 161
DB 240 SEASSEASLPIDEVESMEDVDVGSVSAAPIPK-----KESRKSSNSL 287
QY 162 VEELR 166
DB 288 YESYR 292

RESULT 9
ID Q9P421 PRELIMINARY; PRT; 3839 AA.
AC Q9P421;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE RELATED TO TOM1 PROTEIN.
GN B1122.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Align V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356834; CAB92704.1; -
SQ SEQUENCE 3839 AA; 427187 MW; 6A9E57DEB48A5D9D CRC64;

Query Match 9.8%; Score 91.5; DB 3; Length 3839;
Best Local Similarity 24.6%; Pred. No. 41;
Matches 47; Conservative 29; Mismatches 72; Indels 43; Gaps 9;

QY 3 OSILRHVTSNLSQ-SPAPCRHRTLSKAGSD-----AGDSRSDSPLPHRKDSW 54
DB 2338 KTLSTAVMSQGLPLPASG-----TAGDQAEDEFUSGSSVSD-LTDDRETFDL 2389
QY 55 FYARREASAIL--GLDQKISHLTDELDFDDVOKARAVR--LVEDNEDSDSQSF 110
DB 2390 Y----RNSTLGLMEPRGDFEDEDDEDMYDEQYDDELJYGDMSQDNEDNPSDEED 2446
QY 111 PVKDYMSNGTIDNFSLOTATIMPTGTP-----RSDDGDALFFGDKSKSQASNV 162
DB 2447 DLGEMGEGG-----MPQPGVVEVLNGENDEEDNDMDDEDEDEDE 2494
QY 163 EELRQQAQME 173
DB 2495 QELSEDEDEE 2505

RESULT 10
ID Q9NZ23 PRELIMINARY; PRT; 219 AA.

AC Q9NZ23;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HSPC177.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human full length cDNA cloned from cd34+ stem cells."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF161525; AAF29140.1; -
SQ SEQUENCE 219 AA; 24571 MW; C95695F66F07E126 CRC64;

Query Match 9.7%; Score 91; DB 4; Length 219;
Best Local Similarity 26.4%; Pred. No. 1.6;
Matches 34; Conservative 25; Mismatches 40; Indels 30; Gaps 6;

QY 68 LDOKISHLTDELDFDDVOKAR-----AVR-----RGLVEDNEDSDSQSFPV 112
DB 31 IDKKISRL-DAELVYKDKIKKREGPAKVMYKQKALEVLKQKRYEQORDNLAAQSFNM 89
QY 113 RDKYMSNGTIDNFSLOTATIMPTGTPRSDDDGDALFFGDKSKSQASNVDELROQQAQM 172
DB 90 E---QANYTIQSLK-DTKTV-----DAMKLGKEMKRAYQVKVQKIDQIEDLQDL 135
QY 173 EEALKTAKQ 181
DB 136 EDMEDANE 144

RESULT 11
ID Q17652 PRELIMINARY; PRT; 1707 AA.
AC Q17652;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE C05C12.3 PROTEIN.
GN C05C12.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68333; CAA92726.1; -
DR INTERPRO; IPR002111; -
SQ SEQUENCE 1707 AA; 194652 MW; 2E69AAE8603ED60D CRC64;

Query Match 9.6%; Score 90.5; DB 5; Length 1707;
Best Local Similarity 21.9%; Pred. No. 19;
Matches 48; Conservative 30; Mismatches 62; Indels 79; Gaps 11;

QY 12 NSQLGSPAGRHRTLSKAGSDADGSDRS-----DSPLPHRKDSW 55
DB 1501 NKLKGFISGKYKRGSGFGSGGGSSDNKSLPNNVPMITYDGPSP-IGSRRTSGQ 1559
QY 56 YARREASAILGLDQKISHLTDELDFDDVOKARAVRGLVEDNEDSDSQS----- 108
DB 1560 YLKRDSLQA-----KKKI---TENRRSSL--EQKIPSIQFNLMDQDESAESATEEVS 1610
QY 109 SFPV-----RDK-----YMSNG-----TIDN 124

Db 1611 SIPVQMRVQVTESSDKSLSEDLITREDAPPTSLNPRRRHALYSTIADIETEDD 1670

QY 125 FSLDTAITMPTGTRSDDDGDALEFGDKKSKQDASNDVVE 163
| : : : : : ||| : : : : : |

Db 1671 FYADSPVPMPTVPQADGS--FFGENDSRYQRDSDYE 1707

RESULT 12

Q41042 PRELIMINARY; PRT; 611 AA.

AC Q41042;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE PROTEIN LOCALIZED IN THE NUCLEOLI.

OS Plasm stivum (Garden pea).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

OC Fabales; Fabaceae; Papilionoideae; Pisum.

OX NCBI_TaxID=3888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ALASKA; TISSUE=APICAL MERISTEM;

RA Tong C.G., Hsieh H.L., Blumenthal S., Reichler S., Balk J., Roux S.J.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL: L43510; AAA74208.1; -

DR HSSP: P09651; IUP1.

DR MENDEL: 15263; Pissa; 2406; 15263.

DR INTERPRO: IPR000504; -

DR PFAM: PF00076; rrm; 2.

SQ SEQUENCE 611 AA; 64774 MW; 7CFC749EF8ADEB64 CRC64;

Query Match 9.6%; Score 90; DB 10; Length 611;
Best Local Similarity 23.9%; Pred. NO. 6.4;
Matches 45; Conservative 33; Mismatches 62; Indels 48; Gaps 9;

QY 19 APRGHRDTLSKAGSDGDSRSDPLPHER-----GKSDWFWARREASAILGLQKIS 73
| : : : : : ||| : : : : : |

Db 176 APAR--KAASSDEEDESSEDEEAPKAVVPAVKPA---AAKKKAESS-----DSDSS 226
| : : : : : ||| : : : : : |

QY 74 HLTDDELDA--LPDDYQKARAVRGLVEDNEDS-----DSQSFPPVRDKYMSNG----- 120
| : : : : : ||| : : : : : |

Db 227 EDSDEEDTKPTVTAVSKSAVAKKTEDDEDSSDEEDNKTSPSGNKKPVTV 286
| : : : : : ||| : : : : : |

QY 121 -----TIDNFSLDTAITMPTGTRSDDDGDALEFGDKKSKQDASNDVVEELRQQAQME 173
| : : : : : ||| : : : : : |

Db 287 SKKEDKMNVDKSSD-----SDESDS-----DEESEDEPSKTPQKKTKDVEMI 332
| : : : : : ||| : : : : : |

QY 174 EALTKAQ 181
| : : : : : ||| : : : : : |

Db 333 DADKSKK 340

RESULT 13

Q26189 PRELIMINARY; PRT; 757 AA.

AC Q26189;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE MICRONEME PROTEIN-1 (FRAGMENT).

OS Plasmodium vivax.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5855;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RX MEDLINE=88139712; PubMed=2449454;

RA Cochrane A.H., Gwadz R.W., Kamboj K.K.;

RT "Isolation of three distinct parasite populations expressing different

RT circumsporozoite proteins from the Philippine strain of Plasmodium

RT knowlesi.";

RL J. Clin. Microbiol. 26:178-181(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RX MEDLINE=91004213; PubMed=2170017;

RA Adams J.H., Hudson D.E., Torii M., Ward G.E., Wellem T.E., Alkawa M.,

RA Miller L.H.;

RT "The Duffy receptor family of Plasmodium knowlesi is located within

RT the micronemes of invasive malaria merozoites.";

RL Cell 63:141-153(1990).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RX MEDLINE=91187056; PubMed=1849231;

RA Fang X.D., Kaslow D.C., Adams J.H., Miller L.H.;

RT "Cloning of the Plasmodium vivax Duffy receptor.";

RL Mol. Biochem. Parasitol. 44:125-132(1991).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RX MEDLINE=92357776; PubMed=1496004;

RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;

RT "A family of erythrocyte binding proteins of malaria parasites.";

RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RA Tsuboi T., al-Yaman F., Prickett M.D., Alpers M.P., Adams J.H.;

RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE OF 1-751 FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 32-1;

RX MEDLINE=95048797; PubMed=7960140;

RA Tsuboi T., Kappe S.H., al-Yaman F., Prickett M.D., Alpers M.,

RA Adams J.H.;

RT "Natural variation within the principal adhesion domain of the

RT Plasmodium vivax duffy binding protein.";

RL Infect. Immun. 62:5581-5586(1994).

DR EMBL: L23072; AAA61769.1; -

FT NON_TER 1 1

FT 757 757

SQ SEQUENCE 757 AA; 84006 MW; A3B354E25F212D3C CRC64;

Query Match 9.6%; Score 90; DB 5; Length 757;
Best Local Similarity 25.7%; Pred. NO. 8.2;
Matches 47; Conservative 24; Mismatches 72; Indels 40; Gaps 8;

QY 6 LRHSVTNSQLGSPAPGRHDTLSKAGSDADGDSRSDPLPFRGKSDWFWARREASAI 65
| : : : : : ||| : : : : : |

Db 429 LRHSDNSD---GPAEFAESTKSAESMANPDSNKGTEG--KGQDNMAKATKDSNSS 482
| : : : : : ||| : : : : : |

QY 66 LGLDQKISHLTDDLDALFDDVQKARAVRGLVEDNEDSDSQSFPPVRDKYMSN---GTI 122
| : : : : : ||| : : : : : |

Db 483 DG-----TSSATGDTTDAV-----DREINKGVPE-----RDKTVGSKGGGE 520
| : : : : : ||| : : : : : |

QY 123 DNFSLDTAITMPTGTRSDDDGDALEFGDKKSKQDASNDVVEE---LRQQAQMEELAKTA 179
| : : : : : ||| : : : : : |

Db 521 DNSANKDAATVVGEDRIENSAG-----GSTNDRSKNTEKNGASTPDSKQSEDTALS 574
| : : : : : ||| : : : : : |

QY 180 QKE 182
| : : : : : ||| : : : : : |

Db 575 KTE 577

RESULT 14

Q60465 PRELIMINARY; PRT; 590 AA.

AC Q60465;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE PEST REPEATS ACIDIC PROTEIN GADD34.

GN GADD34.
OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94187707; PubMed=8139541;
RA Zhan Q., Lord K.A., Alamo I. Jr., Hollander M.C., Carrier F., Ron D.,
RA Kohn K.W., Hoffman B., Liebermann D.A., Fornace A.J. Jr.;
RT "The gadd and myd genes define a novel set of mammalian genes encoding
RT acidic proteins that synergistically suppress cell growth.";
RL Mol. Cell. Biol. 14:2361-2371(1994).
DR EMBL: L28147; AAA36983.1;
SQ SEQUENCE 590 AA; 64527 MW; B3D879BDACBAE6D6 CRC64;

Query Match 9.5%; Score 89.5; DB 11; Length 590;
Best Local Similarity 25.7%; Pred. No. 6.7;
Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;

OY 15 LGSPPAGRHD--RTLSKAGSDAGDSRSPLPHFRGKSDWFFARREAAAILGLDQK 71
DB LSPSPSPHDFLKAWYRGEDEDD-----DSWGSAAEEG-----KA 348
OY 72 ISHLTDELDALEFDVQKARAVRGLVEDNEDSDSPFPVRDKYMSNGTIDNFS----- 126
DB LSPSPSPHDFL-----KAWYRPG--EDTED-DQSDWGSAAEK-----DGLAQTFATPHTS 397
OY 127 --LDTAITMPTCPTRSDDDGALFFGDKKSDASNVVDVLRQQAQ 171
DB AFLKTVWCPCGEDTEDDCEVVV-----PEDSEADPKSPSHDAQ 438

RESULT 15
Q26188 ID Q26188 PRELIMINARY; PRT; 750 AA.
AC Q26188;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE MICRONEME PROTEIN-1 (FRAGMENT).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PAPUA NEW GUINEA 29-1;
RX MEDLINE=88139712; PubMed=2449454;
RA Cochrane A.H., Gwadz R.W., Kamboj K.K.;
RT "Isolation of three distinct parasite populations expressing different
RT circumsporozoite proteins from the Philippine strain of Plasmodium
RT knowlesi.";
RL J. Clin. Microbiol. 26:178-181(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=PAPUA NEW GUINEA 29-1;
RX MEDLINE=91004213; PubMed=2170017;
RA Adams J.H., Hudson D.E., Torii M., Ward G.E., Welles T.E., Alkawa M.,
RA Miller L.H.;
RT "The Duffy receptor family of Plasmodium knowlesi is located within
RT the micronemes of invasive malaria merozoites.";
RL Cell 63:141-153(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=PAPUA NEW GUINEA 29-1;
RX MEDLINE=91187056; PubMed=1849231;
RA Fang X.D., Kaslow D.C., Adams J.H., Miller L.H.;
RT "Cloning of the Plasmodium vivax Duffy receptor.";
RL Mol. Biochem. Parasitol. 44:125-132(1991).
RN [4]
RP SEQUENCE FROM N.A.

RC STRAIN=PAPUA NEW GUINEA 29-1;
RX MEDLINE=92357776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
OC Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=PAPUA NEW GUINEA 29-1;
RA Tsuboi T., al-Yaman F., Prickett M.D., Alpers M.P., Adams J.H.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=PAPUA NEW GUINEA 29-1;
RX MEDLINE=95048797; PubMed=7960140;
RA Tsuboi T., Kappe S.H., al-Yaman F., Prickett M.D., Alpers M.,
RA Adams J.H.;
RT "Natural variation within the principal adhesion domain of the
RT Plasmodium vivax duffy binding protein.";
RL Infect. Immun. 62:5581-5586(1994).
DR EMBL: L23071; AAA61768.1;
FT NON_TER 1
FT NON_TER 750
FT NON_TER 750
SQ SEQUENCE 750 AA; 83303 MW; DFF3A2270ECFBBB1 CRC64;

Query Match 9.5%; Score 89.5; DB 5; Length 750;
Best Local Similarity 24.0%; Pred. No. 8.9;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

OY 6 LRHSVTNSQLSGPAPGRHRTLSKAGSDAGDSRSPLPHFRGKSDWFFARREAAAI 65
DB LRHSKNSDSDGPA-----ESMANPDSNRKGTG-----KGQNDMAKATKDSNSS 475
OY 66 LGLDQKISHLTDELDALEFDVQKARAVRGLVEDNEDSDSPFPVRDKYMSN---GTI 122
DB LGLDQKISHLTDELDALEFDVQKARAVRGLVEDNEDSDSPFPVRDKYMSN---GTI 122
OY 476 DG---TSSATGDTTDAV-----DREINKGVPE-----RDKTVGSKDGGGE 513
OY 123 DNFSLDTAITMPTCPTRSDDDGALFFGDKKSDASNVVDVEE---LRQQQAQMEALKTA 179
DB DNFSLDTAITMPTCPTRSDDDGALFFGDKKSDASNVVDVEE---LRQQQAQMEALKTA 179
OY 514 DNSANKDAATVVGEDRIENSAG-----GSTNDRSRNDTEKNGASTPDSKQSEDAVALS 567
OY 180 KOE 182
DB 568 KTE 570

Search completed: April 25, 2001, 10:18:19
Job time: 325 sec

